Sketches of an Elephant
A Topos Theory Compendium

VOLUME 1

PETER T. JOHNSTONE

Reader in the Foundations of Mathematics
University of Cambridge

CLARENDON PRESS • OXFORD

2002

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

with an associated company in Berlin

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2002

A catalogue record of this title is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 853425 6 (Vol. 1)

ISBN 0 19 851598 7 (Vol. 2)

ISBN 0 19 852496 X (2 Vol. set)

10 9 8 7 6 5 4 3 2 1

Typeset by
Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
T. J. International Ltd, Padstow, Cornwall

PREFACE

Introduction and apologia

Four men, who had been blind from birth, wanted to know what an elephant was like; so they asked an elephant-driver for information. He led them to an elephant, and invited them to examine it; so one man felt the elephant's leg, another its trunk, another its tail and the fourth its ear. Then they attempted to describe the elephant to one another. The first man said 'The elephant is like a tree'. 'No,' said the second, 'the elephant is like a snake'. 'Nonsense!' said the third, 'the elephant is like a broom'. 'You are all wrong,' said the fourth, 'the elephant is like a fan'. And so they went on arguing amongst themselves, while the elephant stood watching them quietly.[1]

The notion of topos resembles an elephant in that it is possible to come up with very different descriptions of what topos theory is about, depending on the direction from which you approach the subject. It was André Joyal, in his lecture at the 1981 Cambridge Summer Meeting on Category Theory, who first drew attention to the parallel, and I am grateful to him for allowing me to use it as the basis for the title of the present work. He listed seven different descriptions of 'what a topos is like':

(i) 'A topos is a category of sheaves on a site'

(ii) 'A topos is a category with finite limits and power-objects'

(iii) 'A topos is (the embodiment of) an intuitionistic higher-order theory'

(iv) 'A topos is (the extensional essence of) a first-order (infinitary) geometric theory'

(v) 'A topos is a totally cocomplete object in the meta-2-category $\mathfrak{CART}$ of cartesian (i.e., finitely complete) categories'

(vi) 'A topos is a generalized space'

(vii) 'A topos is a semantics for intuitionistic formal systems'

In the 20 years that have passed since that lecture, the category-theory community has added a few more descriptions to the list (and the theoretical computer scientists have contributed yet others); for example,

(viii) 'A topos is a Morita equivalence class of continuous groupoids'

(ix) 'A topos is the category of maps of a power allegory'

[1] The story of the blind men and the elephant exists in numerous versions in Indian folklore. This version is freely adapted from E. J. Robinson, *Tales and Poems of South India* (London, 1885).

(x) 'A topos is a category whose canonical indexing over itself is complete and well-powered'

(xi) 'A topos is the spatial manifestation of a Giraud frame'

(xii) 'A topos is a setting for synthetic differential geometry'

(xiii) 'A topos is a setting for synthetic domain theory',

and so on. But the important thing about the elephant is that 'however you approach it, it is still the same animal'; this book is an attempt to demonstrate that the same is true of topos theory.

A brief bibliography of topos theory

This book resembles an elephant in another respect: namely its gestation period. To explain (and apologize for) this, I have to go back still further into the history of the subject. I must also crave the reader's indulgence for what may seem an excessively egocentric view of that history in the paragraphs which follow.

The original notion of topos, as a 'generalized space' suitable for supporting the exotic cohomology theories required in algebraic geometry, sprang from the fertile brain of Alexandre Grothendieck in the early 1960s, and was developed in his 'Séminaire de Géometrie Algébrique du Bois-Marie' particularly during the academic year 1963–64. The duplicated notes of that seminar circulated widely among algebraic geometers and category-theorists over the next decade, until Springer-Verlag did the world a service by publishing a revised and expanded edition in three volumes of *Lecture Notes in Mathematics* in 1972 [36]. But by then the subject had already been 'reborn' in its second incarnation, as an elementary theory having links with higher-order intuitionistic logic, through the collaboration of Bill Lawvere and Myles Tierney during 1969–70 (which, in turn, built upon Lawvere's work on providing a categorical foundation for mathematics, which had been developing since the early 1960s).

As a graduate student in the early 1970s, I was fortunate to 'get in on the ground floor' of elementary topos theory, by attending Tierney's summer school lectures [1167] at Varenna in 1971, and later Saunders Mac Lane's lectures during his Cambridge sabbatical in autumn 1972. At that time the only general accounts of the subject in print were the lecture notes [649] by Anders Kock and Gavin Wraith, which were not all that widely available, and Peter Freyd's survey article [371], whose view of the subject was idiosyncratic to say the least – though they were joined over the next two years by the notes of Tierney's lectures, already cited, and by Wraith's second account [1236].

More or less immediately after completing my Ph.D. in 1974, I embarked on the writing of what became the first account in book form of the elementary theory of toposes, published by Academic Press in 1977 [504]. With the brash self-confidence of youth, I attempted to provide between one pair of covers a full account of as many aspects of the elephant as I could muster: the result was

a book which, though universally acknowledged as 'far too hard to read, and not for the faint-hearted',[2] remained in demand for a long time – and continues to be widely cited even today, despite being nearly 25 years out of date.

For some time [504] was the only available book on elementary topos theory. But by the time of Joyal's 1981 lecture it had been joined by a second, Robert Goldblatt's *Topoi: the categorial analysis of logic* [411], and over the next decade or so a good many other books appeared which either were devoted to topos theory or contained a good deal of topos-theoretic material: *Teoria toposurilor* [997] by Alexandru Radu in 1981–2, *Toposes, triples and theories* [87] by Michael Barr and Charles Wells in 1985, *Introduction to higher order categorical logic* [682] by Joachim Lambek and Philip Scott in 1986, *Toposes and local set theories* [95] by John Bell in 1988, *Categories, allegories* [381] by Peter Freyd and Andre Scedrov in 1990, *Models for smooth infinitesimal analysis* [856] by Ieke Moerdijk and Gonzalo Reyes in 1991, *Lecture notes on topoi and quasitopoi* [1248] by Oswald Wyler, also in 1991, *Sheaves in geometry and logic* [751] by Saunders Mac Lane and Ieke Moerdijk in 1992, *Elementary categories, elementary toposes* [760] by Colin McLarty, also in 1992, *Relative category theory and geometric morphisms* [238] by Jonathan Chapman and Frederick Rowbottom, yet again in 1992, and finally the third volume of Francis Borceux' remarkable *magnum opus Handbook of categorical algebra* [147] in 1994.

Each of these books contributed a new view of the elephant, illuminating features not visible from other angles, and thus each of them was valuable. But I hope I shall not offend the authors listed above by saying that none of them, with the possible exception of the last, really attempted to give a picture of the whole animal: indeed, in several cases the concentration on one part of its anatomy is self-evident from the title.

Meanwhile, my own book had gone out of print in 1987, despite a corrected reprint (some time after the original stock had run out) in 1985. There was some suggestion at that time that Academic Press might be willing to publish a revised second edition of the book, if I were willing to write it; but their procrastination over, and lack of enthusiasm for, the 1985 reprint did not fill me with enthusiasm for that course of action. Another reason for my reluctance was the extent to which the subject had grown since the mid-1970s: since the principal merit of my book, in comparison with the others by then available or forthcoming, was its comprehensiveness, there would be no point in replacing it by anything other than a comprehensive account of the subject as it stood by then, and such an account would inevitably be very much longer than the 1977 book.

[2] The quotation is from the comments of an anonymous referee consulted by OUP about the present book.

The three blind men

By the mid-1980s Ieke Moerdijk had also become concerned (despite his involvement as co-author in two of the books mentioned above) about the lack of any single account of topos theory which could properly claim to give a comprehensive picture of the subject. However, like me he felt daunted by the prospect of writing such a book single-handed. Sometime early in 1986, he mentioned to me his idea that what was needed was a 'handbook' type publication in which a small team of experts would contribute extended essays on the different aspects of the subject, under the overall supervision of a co-ordinating editor – he felt that I was the ideal person to take on the latter rôle. I was not, at first, greatly taken with this idea: as I saw it, the chief problem with the 'handbook' style was that, although the elephant has many different aspects, their interconnectedness is such that it is impossible fully to appreciate any one of them without making some reference to most of the others. So I foresaw that the job of a co-ordinating editor, for such a volume, would be more than usually nightmarish.

Nevertheless, discussions between Ieke and myself about the idea continued to take place on the odd occasions when we met, and they soon came to involve my former student Andrew Pitts as another possible author. The 'handbook' idea gradually metamorphosed into something more like a 'Bourbaki' volume or the *Compendium of continuous lattices* [404], in which each author would have individual responsibility for the first drafts of particular sections, but the final text would have to be agreed by all the authors (and the individual authorship of the different sections would not be publicly identified). At a meeting between the three of us in December 1987, we formally agreed to go ahead with a three-author work on that basis, and a rough table of contents – including the division of the book into six parts, although they did not correspond exactly to the six parts in the present table of contents – was worked out, together with the responsibilities of the three of us for drafting the more introductory sections. The working title *Sketches of an elephant* was also agreed at this meeting.

By the late summer of 1988 we had enough material in draft to approach publishers with a definite proposal, and we signed an agreement with Oxford University Press in January 1989. At the time (I blush to recall it now) we hoped that we would be able to deliver a complete typescript to the publishers by the end of September 1991. We also estimated that the overall length of the book might be around 1000 pages; in the event, that estimate has proved more accurate than the temporal one.

Over the next couple of years, a substantial quantity of draft material was written, with P. T. J. working mainly on Parts A and B, I. M. on Part C and A. M. P. on Part D. Ieke produced large quantities of draft material, principally on topics from Chapters C2 and C3, although virtually none of it has survived into the version that exists today; on the other hand, stylistic experts may just be able to detect the origins of the present Sections D1.1–D1.3 in a draft written by Andrew, even though it has been heavily rewritten by me. But it had

become clear by early 1991 that the project was a much larger one than we had realized: in particular, the problems of co-ordination between the three of us, and of reaching agreement on a text that all three could be happy with, seemed overwhelming. It was at least partly these problems that led Ieke and Andrew independently to decide to withdraw from the project in late 1992.

I should make it clear that, despite their not seeing the project through to completion, I remain enormously indebted to both Ieke and Andrew. If it had not been for their support in the early stages, I should never have dared to embark on the writing of a work on this scale, and their encouragement kept me going until the book had progressed far enough to acquire its own momentum. Also, their constructive criticisms of my early drafts of Part A caught many infelicities that I should probably not have noticed myself. And on a more mundane level, Andrew taught me to use LaTeX,[3] without which the project would have been not just foolhardy but downright impossible.

One blind man

I thus found myself in early 1993 faced with three options: to abandon the project myself, to try to find other co-authors, or to accept the burden of writing the whole book myself. The first of these was unattractive because it would mean abandoning the large amount of material (including a complete draft of Part A) that I had already written, and the second did not appeal since I could not think of any potential co-author who would be as easy to work with, or as productive, as I knew Ieke and Andrew to be. So I decided to go ahead with the third option, although I knew that that would involve a 'long hard slog' over several years – I found it impossible to estimate how many.

With my other commitments to teaching, research and administration, progress was indeed slow: often, several months would pass without my finding time to open the files where the text of the book was stored. In addition, I often felt like Heracles wrestling with the Hydra: every attempt to draft a new section seemed to entail the rewriting of at least two existing ones, so that I appeared to come no nearer to the goal of a complete text. (There was also, at times, the fear that the subject was advancing faster than I could write it down.)

Nevertheless, progress did sometimes get made. One thing which helped to spur the drafting, or revision, of certain sections was the chance to lecture about the material in them, to audiences around the world, on occasions which included the following:

- In June 1994 I was invited to lecture on classifying toposes at Shaanxi Teachers' University, Xi'an, China, through the good offices of Wang Guo-Jun

[3] Whilst I am on that subject, I should also record my indebtedness to another of my former students, Paul Taylor, whose TeX diagram-drawing macros are used throughout the book.

and my former student Zhao Dongsheng. This involved the polishing of material from Chapters A1, D1 and D3.

- In May 1995, Francis Borceux invited me to lecture on locales at a summer school on 'Topological Applications of Category Theory' in Alle-sur-Semois, Belgium: this involved material from Chapters C1 and C3.
- During my sabbatical visit (arranged by Max Kelly) to Sydney, Australia, during February and March 1996, I gave a number of talks in the Sydney Category Seminar covering material from Chapters B1 and B3.
- In July 1997, as an invited speaker at the 'Workshop on Logic, Langage, Information and Computation' (commonly known as WoLLIC) in Fortaleza, Brazil, I had the opportunity (thanks to Ruy de Queiroz) to give a tutorial on categorical logic before the main conference: once again, this provided an occasion to polish the material in Chapters D1 and D2.

I mention these meetings in order publicly to record my thanks to the organizers named: though they were not aware of it, they helped to bring the project closer to completion.

By the beginning of the third millennium, however, it had become clear to me that what was absolutely essential, if the writing of Parts A–D was ever to be finished, was to find an extended period of time in which I could set all other work aside and concentrate full-time on the book. (Although I had at various times begun work on several sections from Part F, I had by then taken the decision that the only sensible route to completion was to concentrate on the first four parts, which were too strongly interconnected to be written separately, and to leave the last two – which, dealing as they do with applications, are less strongly tied to the others – for completion at a later date and publication as a separate volume.) By great good fortune, such an opportunity arose with the invitation to participate in the 'Mathematical Logic Year' at the Institut Mittag-Leffler in Djursholm, Sweden, for a two-month period in January–March 2001. My gratitude to the organizers of that programme (Dag Normann, Viggo Stoltenberg-Hansen and Jouko Väänänen), for providing me with this opportunity, is impossible to exaggerate; and I am also deeply grateful to Kjell-Ove Widman and the staff of the Institut Mittag-Leffler for providing such a near-perfect working environment.

The speed with which everything fell into place, once I had the opportunity for uninterrupted work on the book, greatly surprised me. Although the proportion of the total text which was written in Sweden is not so great, the number of sections which reached (more or less) their final form during my time there is very large: about half of those in Part B, almost all in Part C apart from Chapter C1, and most of those in Chapters D4 and D5. (In addition, a good deal of material was added to the supposedly complete draft of Part A.) Within a month of my return to Cambridge, I was able to produce a complete text of the first four parts, except for three sections in Part D. Copies of this draft were circulated to a number of colleagues (including Ieke Moerdijk and Andrew Pitts,

as well as Martin Hyland, Anders Kock and Gavin Wraith) for their comments and suggestions; the latter (which were, in many cases, of considerable value) were taken into account in the final revision of the text which took place in August–September 2001, after the writing of the last three sections had been completed.

At the time of writing these words, Parts E and F of the original plan remain to be written. In view of my experience so far, I am making no promises about how long it will take to write them; but I hope that interested readers will not have to wait as long for the third volume as they have done for the first two.

How to read this book

'Begin at the beginning,' the King said gravely, 'and go on till you come to the end; then stop.'[4]

Whilst the King of Hearts's advice to the White Rabbit is generally sound, it is not particularly helpful in relation to this book. At the December 1987 meeting to which I have referred, the three of us took a number of decisions about the overall structure and content of the book: despite the many changes of detail which have taken place since then, the general principles which we agreed have by and large survived unchanged. Since some understanding of these principles is important for anyone who seeks to obtain information from this book, I shall describe them here.

The first and most important is that *the book is not a textbook*. It is not addressed to those who are trying to learn about topos theory for the first time,[5] but rather to those who already have some acquaintance with the subject and who wish to deepen their understanding, or to learn about aspects of it which they have not previously encountered. In keeping with this, there are no 'exercises for the reader' (which are often the lazy author's way of leaving out the proofs of results that he can't be bothered to write out in full): instead, all results needed in the book (with the exception of a few results from other areas of mathematics, for which references to standard works are given) are proved in full, and all examples are fully worked out.[6] This is not to say that the book *cannot* be used as a course text: it would, for example, be possible to base an introductory course on the purely category-theoretic side of topos theory on Chapters A1, A2 and A4, although the instructor would have to be fairly selective about the specialized material which has been inserted in those chapters for future reference, and would also probably want to leaven the presentation with some material from Parts C and/or D.

[4] Lewis Carroll, *Alice's Adventures in Wonderland* (London, 1865).

[5] For those who do seek an introduction to the subject, the one by Mac Lane and Moerdijk [751] remains (in the present writer's opinion) the most recommendable.

[6] It should be understood that this policy does not debar the author from leaving routine 'diagram-chasing' and other calculations to the reader.

Secondly, *the book is not intended to be read sequentially.* The whole point of the six-part structure is that the book is seeking to present several different approaches to topos theory, so where one starts reading will depend on the line of approach one wishes to follow. Most readers, however experienced, will probably wish to start by at least skimming through Chapter A1, in order to familiarize themselves with the book's conventions on categorical notation and terminology (about which I shall have more to say below); but thereafter a reader who already knows the basic category theory of toposes may well decide to jump to Chapter B1, C1 or D1, depending on the approach he wishes to take. However, almost any reader will have to be prepared to refer backwards and/or forwards a good deal. Although it would have been theoretically possible to write entirely independent accounts of the geometric and logical aspects of topos theory in Parts C and D (and that was more or less our original intention), in fact the two are so thoroughly interconnected that such a course would have involved an immense amount of duplication (and of course increased the overall length of the book).

There is some duplication, which is quite deliberate: for example, the treatment of natural number objects in Section A2.5 overlaps with that in Section D5.1, the construction of classifying toposes is treated in two different ways in Sections B4.2 and D3.1, and exponentiability of toposes is treated in Sections B4.3 and C4.4. And there are numerous smaller instances, where an individual result is proved in two different ways, using material from different areas of the book. But in general, if there is an optimal proof of a particular theorem, and it clearly belongs (say) in Part D, there does not seem to be any point in writing it out a second time in Part C if it happens to be needed there. So the reader who wishes to see the full picture must be prepared to skip around from one part to another, when necessary. (On the other hand, a reader who is willing to 'take on trust' the results quoted from other parts of the book should be able to follow any given chapter as a (more or less) connected narrative.)

I have at least done my best to make the cross-references easy to use. All numbered references (definitions, lemmas, theorems, examples, etc.) within a given section are numbered in a single sequence: thus 'B3.4.5' denotes the fifth numbered reference (in fact a corollary) in the fourth section of the third chapter of Part B. For references within a given Part, the letter is omitted: thus if the corollary just mentioned is cited elsewhere in Part B, it appears as '3.4.5'. If a lemma, proposition, theorem, corollary or scholium[7] is subdivided into parts, they are normally labelled (i), (ii), (iii) and so on, whereas definitions, examples and remarks are subdivided (*a*), (*b*), (*c*),.... The symbol □ is used to denote the end of a proof; if it appears at the end of the statement of a result, it means

[7] I should perhaps explain that I use the term 'scholium' (literally, a marginal note) to denote something which follows directly from the proof of a preceding result, as opposed to a corollary which follows directly from the statement of the preceding result.

either that no proof is given, or that the proof is contained in the discussion which precedes the statement.

At the December 1987 meeting we adopted the principle that, for every major result to be included in the book, we should seek to find the optimal proof,[8] although that should not preclude us from giving alternative proofs if they helped to shed additional light on the result. (I hope it is unnecessary to add that I am conscious of having not always succeded in this aim.) Partly for this reason, we decided that we would not attempt to give detailed attributions for all the results in the book, since the optimal proof is often not the original one, but has been arrived at through successive polishing of the latter by several hands. Nevertheless, I have attempted to compile a comprehensive bibliography of topos theory, and to give attributions where credit is clearly due to a single source. (I considered attempting to classify the bibliography entries by subject area, as was done in [404], for example; but the task seemed too large and ultimately of doubtful value.) And in most sections I have provided a list of references at the end of the section for those who wish to pursue the topic(s) covered in the section in greater detail.

We also recognized that the choice of terminology and notation, in a book of this nature, was of immense importance. It is impossible to be entirely consistent about notation, in a subject of this size: there simply aren't enough different letters and typefaces for one to avoid re-using some of them in mutually contradictory ways. But I have done my best to ensure local consistency of notation, within each chapter. As regards terminology, I have consciously attempted to use this book to disparage some of the sloppier pieces of inappropriate usage which are common among category-theorists; whether it will have any influence in this respect of course remains to be seen, but readers who are familiar with the existing terminology may well find some of my choices rather surprising at first. I have, however, listed alternatives to my chosen terminology in the index, where they appear in italics, and in most cases my reasons for discarding them will be found in the text at the points to which the index entries refer.

A comprehensive index is also of great importance in a book of this nature. I have done my best to provide one: in addition to listing alternative terminology in italics, as already mentioned, it also contains entries for many of the principal theorems, particularly those which occur in several different versions in different places, and I hope that this will assist the specialist in finding his way quickly to the information which he wants. One point which should be mentioned about the index is that, where an index entry begins with a mathematical symbol, it is listed in the alphabetical order as if the name of the symbol had been spelled out: thus, for example, 'σ-pretopos' will be found between 'sifted coverage' and 'signature'. In addition to the general index, there is also an index of standard notation, ordered alphabetically on the same principle.

[8] That is, the one which the late Paul Erdős would have described as belonging to 'The Book'.

Conclusion

I have already mentioned my indebtedness to several colleagues who have played particular rôles in the story which led up to the writing of this book. But there are many other workers in topos theory to whom I am indebted, for their ideas which have contributed to my own understanding of topos theory, and which have in many cases found their way into the text which follows without specific acknowledgement. It would be quite impossible to thank them all individually, so I hope that they will accept these few words of collective thanks.

Finally, I must thank the staff of Oxford University Press, and in particular Elizabeth Johnston and Alison Jones: not only for their enormous patience in waiting for this text ten years beyond the date when we originally promised to deliver it, but also for their enthusiastic and helpful response when I was finally able to let them know that its delivery was within sight. It has been a pleasure to work with them, and I look forward to doing so again when Parts E and F are ready for publication.

Cambridge
September 2001 P. T. J.

CONTENTS VOLUME 1

CONTENTS VOLUME 2

CONTENTS VOLUME 3*

* Volume 3 is being written and therefore the table of contents is liable to change

PART A

TOPOSES AS CATEGORIES

A1
REGULAR AND CARTESIAN CLOSED CATEGORIES

A1.1 Preliminary assumptions

In this book, we start from the assumption that the reader is already familiar with the basics of category theory: specifically, we assume that he has already seen the definitions of the terms *category*, *functor*, *natural transformation*, *adjunction*, *limit* and *monad*. These definitions can, of course, be found in any of the standard textbooks on category theory (for example, [742], [381] or [147]). However, there are a number of conventions on which the standard texts differ, and we need to state our position on these before proceeding further.

First, we (generally) write functions on the left of their arguments, and therefore compose morphisms in a category 'from right to left': fg means 'g followed by f'. (We normally denote composition, as here, simply by juxtaposition; but we shall occasionally write $f \cdot g$ or $f \circ g$ in order to improve the readability of complicated expressions.) Next, we call a functor *faithful* if it is 'injective on hom-sets': that is, $F: \mathcal{C} \to \mathcal{D}$ is faithful iff, given morphisms f and g of $\mathcal{C}$, the equations $\mathrm{dom}\, f = \mathrm{dom}\, g$, $\mathrm{cod}\, f = \mathrm{cod}\, g$ and $F(f) = F(g)$ together imply $f = g$. We do not require that F should also reflect isomorphisms; if it does so (and is faithful), we call it *conservative*. (Of course, there are circumstances in which any faithful functor is automatically conservative: for example, if the domain category $\mathcal{C}$ is *balanced* (i.e. has the property that every morphism which is both monic and epic is an isomorphism), or if F is full as well as faithful.)

A *full* subcategory, of course, is one whose inclusion functor is full; but when dealing with full subcategories we shall generally assume (sometimes without saying so explicitly) that they are also *replete*, i.e. that any object of the ambient category isomorphic to one in the subcategory is itself in the subcategory. Thus full subcategories of $\mathcal{C}$ correspond to classes of objects of $\mathcal{C}$ which are closed under isomorphism. In particular, for us a *reflective* subcategory will always mean a full, replete subcategory whose inclusion functor has a left adjoint. We use the term *reflection* for an adjunction whose right adjoint is full and faithful, and *reflector* for a monad which is idempotent (i.e. one whose multiplication is an isomorphism); it is well known that these three concepts are essentially the same. The following, related, result seems not to be widely known, however; and since we shall need it occasionally, we sketch its proof here.

Lemma 1.1.1 *Let $F: \mathcal{C} \to \mathcal{D}$ be a functor having a right adjoint G. If there is any natural isomorphism (not necessarily the counit of the adjunction) between FG and the identity functor on $\mathcal{D}$, then $(F \dashv G)$ is a reflection.*

Proof Using the given isomorphism, we may transfer the comonad structure on FG arising from the adjunction to a comonad structure (ϵ, δ) on $1_{\mathcal{D}}$. But the monoid of natural endomorphisms of the identity functor on a category is always commutative; so, since δ is a one-sided inverse for ϵ, we deduce that they are inverse isomorphisms. Transferring back again, we see that the counit of $(F \dashv G)$ is an isomorphism, which is well known to be equivalent to the adjunction being a reflection. □

By an *equivalence* between categories $\mathcal{C}$ and $\mathcal{D}$, we mean a pair of functors $F: \mathcal{C} \to \mathcal{D}$, $G: \mathcal{D} \to \mathcal{C}$ equipped with (a choice of) natural isomorphisms between the composites FG and GF and the respective identity functors. It is well known that, if such isomorphisms exist, they may be chosen to be the unit and counit of an adjunction (in either direction) between F and G; and we shall tacitly assume that this has been done whenever we need it. If F is one half of an equivalence of categories, then it is not only full and faithful but also *essentially surjective* on objects, that is, every object of its codomain $\mathcal{D}$ is isomorphic to one in the image of F. Conversely, if F has these three properties, and we are able to choose for each object B of $\mathcal{D}$ an object GB of $\mathcal{C}$ and an isomorphism $\alpha_B: FGB \to B$, then G extends to a functor $\mathcal{D} \to \mathcal{C}$ such that F and G form an equivalence of categories (α being one of the two natural isomorphisms required). We shall call F a *weak equivalence* if it is full, faithful and essentially surjective on objects.

The terminology just introduced is in keeping with a general principle which will be present throughout this book, though not always explicitly stated: namely, that we should make the smallest possible demands on the 'metatheory' within which we interpret the theory of categories (and in particular we shall not assume that it satisfies any form of the axiom of choice, so that there is a distinction to be made between the assertion that something exists in all cases, and the assertion that we are given a function assigning to each case a choice of that thing). This is because we shall need, at various points, to re-interpret the results of elementary category theory in several different metatheories, and we do not want to have to go through the chore of re-proving these results in each new context. (However, for the present, readers who feel uncomfortable about this may interpret everything we say as applying to categories defined within some fixed universe of sets.)

Thus, when we say '$\mathcal{C}$ is a category', what we mean is that we have given interpretations (in some suitable metatheory) of the statements 'A is an object of $\mathcal{C}$' and 'f is a morphism of $\mathcal{C}$', together with interpretations of the notions 'domain', 'codomain', 'identity morphism' and 'composite' which collectively satisfy the axioms of the first-order theory of categories. We do not assume that the objects or the morphisms of $\mathcal{C}$ form sets (in any formalized sense); nor even that, given a pair of objects A and B of $\mathcal{C}$, the morphisms $A \to B$ in $\mathcal{C}$ form a

set. Whenever (in a context where we are given a particular model of set theory) we wish to make the first or the second of these assumptions, we shall draw attention to it by calling $\mathcal{C}$ a *small category* or a *locally small category*, as the case may be.

As mentioned above, we assume familiarity with the notion of *monad*; in particular, we assume that the reader has seen the constructions of the *Eilenberg–Moore category* $\mathcal{C}^{\mathbb{T}}$ and of the *Kleisli category* associated with a monad $\mathbb{T}$ on $\mathcal{C}$. We shall have occasion to use the Crude Monadicity Theorem (sometimes known as the Crude Tripleability Theorem) in its 'reflexive-coequalizer' form; recall that a parallel pair of morphisms $f, g \colon A \rightrightarrows B$ is called *reflexive* if there exists $s \colon B \to A$ with $fs = gs = 1_B$.

Theorem 1.1.2 *Let $U : \mathcal{D} \to \mathcal{C}$ be a functor, and suppose*

(i) *U has a left adjoint,*

(ii) *U is conservative, and*

(iii) *$\mathcal{D}$ has and U preserves coequalizers of reflexive pairs.*

Then U is monadic, *i.e. the comparison functor $K : \mathcal{D} \to \mathcal{C}^{\mathbb{T}}$ (where $\mathbb{T}$ is the monad on $\mathcal{C}$ induced by U and its left adjoint) is part of an equivalence of categories.*

Proof Most of the standard texts on category theory give this theorem without the restriction to reflexive pairs in condition (iii); but an examination of the proof will show that only coequalizers of reflexive pairs are used in it. □

At various points, we shall need to appeal to the standard 'adjoint lifting theorems' for functors between categories of algebras. For convenience, we state them here:

Proposition 1.1.3 *Suppose given a diagram of categories and functors*

$$\begin{tikzcd}
\mathcal{D} \arrow[r, "\overline{T}"] \arrow[d, "U"] & \mathcal{D}' \arrow[d, "U'"] \\
\mathcal{C} \arrow[r, "T"] & \mathcal{C}'
\end{tikzcd}$$

such that U and U' are monadic and the square commutes up to natural isomorphism.

(i) *If $\mathcal{D}$ has coequalizers of reflexive pairs and T has a left adjoint, then $\overline{T}$ has a left adjoint.*

(ii) *If T and $\overline{T}$ also commute with the left adjoints F and F' of U and U' (more specifically, if the canonical natural transformation $F'T \to \overline{T}F$ which is the*

'mate' of the isomorphism $TU \cong U'\overline{T}$ is an isomorphism) and T has a right adjoint, then $\overline{T}$ has a right adjoint.

Proof See [503], for example. □

We shall occasionally wish to consider 'T-algebras' when T is a mere endofunctor of a category $\mathcal{C}$; that is, pairs (A, α) where $A \in \text{ob}\,\mathcal{C}$ and $\alpha\colon TA \to A$. (The context – and the typeface – will usually make it clear whether we are talking about 'mere' T-algebras in this sense, or about $\mathbb{T}$-algebras in the sense of Eilenberg and Moore.) We recall the following useful fact about T-algebras, first observed by J. Lambek [665]:

Lemma 1.1.4 *Let $T\colon \mathcal{C} \to \mathcal{C}$ be a functor, and suppose the category $\mathcal{C}^T$ of T-algebras has an initial object (A, α). Then α is an isomorphism.*

Proof $(TA, T\alpha)$ is also a T-algebra, and so we have a unique morphism $f\colon A \to TA$ such that the left-hand square in

$$\begin{tikzcd}
TA \arrow[r, "Tf"] \arrow[d, "\alpha"] & TTA \arrow[r, "T\alpha"] \arrow[d, "T\alpha"] & TA \arrow[d, "\alpha"] \\
A \arrow[r, "f"] & TA \arrow[r, "\alpha"] & A
\end{tikzcd}$$

commutes. But the right-hand square trivially also commutes, so the composite αf (is a T-algebra homomorphism, and hence) must be the identity on A. Now the commutativity of the left-hand square yields $f\alpha = T(\alpha f) = 1_{TA}$; so f is a two-sided inverse for α. □

In the definition of a category it is normally insisted that 'hom-sets are disjoint', i.e. (in elementary terms) that each morphism of a category has a uniquely specified domain and codomain. The actual practice of mathematics frequently falls short of this ideal (even in defining the category of sets, one has to adopt an 'unnatural' definition of function), and in order to deal with this we shall find it convenient to borrow the notion of *protocategory* introduced by Freyd and Scedrov [381]. A protocategory $\mathcal{C}$ consists of two types of entities, objects and protomorphisms; the latter may be thought of as possible names for morphisms in the category which $\mathcal{C}$ defines. These data come equipped with two primitive predicates, a *source–target predicate* relating a protomorphism f and a pair of objects A, B (written '$f\colon A \to B$' and read 'f is a name for a morphism from A to B') and a *composition predicate* relating three protomorphisms f, g, h (written '$fg = h$' and read 'h is a possible composite of f and g'; this second predicate is usually, but not necessarily, functional in the sense that, given f and g, there is at most one h satisfying $fg = h$). The key axiom states that, if $g\colon A \to B$ and $f\colon B \to C$, there is exactly one h satisfying $h\colon A \to C$ and $fg = h$.

Of course, we also need axioms to express associativity of composition and existence of identities; but we do not need to write these out in detail, since they may be inferred from the information that the category *generated by* a protocategory $\mathcal{C}$ has the same objects as $\mathcal{C}$ and the instances of the source–target predicate (i.e. the triples (f, A, B) satisfying $f: A \to B$) as morphisms. We shall normally use the same symbol $\mathcal{C}$ for a protocategory and the category which it generates; since we shall only use protocategories as a quick way of specifying particular categories which we want to study, we do not need to develop an independent theory of protocategories.

An example will help to clarify the ideas introduced in the last paragraph.

Example 1.1.5 Given a category $\mathcal{B}$, a category $\mathcal{C}$ is said to be *structured over* $\mathcal{B}$ if it can be generated by a protocategory whose objects come equipped with an *underlying-object map* $A \mapsto |A|$ to the objects of $\mathcal{B}$, whose protomorphisms are the morphisms of $\mathcal{B}$, whose source–target predicate is contained in the relation 'dom $f = |A|$ and cod $f = |B|$' (and has the property that $1_{|A|}: A \to A$ for all objects A), and whose composition predicate is the composition in $\mathcal{B}$. $\mathcal{C}$ is thus specified completely by giving its objects, the underlying-object map and the source–target predicate. For example, the category **Gp** of groups is structured over the category **Set** of sets: its objects are all groups, $|A|$ is the (usual) underlying set of a group A and the source–target predicate is the relation '... is a homomorphism from ... to ...'. (Similarly for other categories of universal algebras, for categories of topological or uniform spaces, etc.; and for the category $\mathbf{Gp}(\mathcal{C})$ of group objects in a category $\mathcal{C}$ with finite products, regarded as a structured category over $\mathcal{C}$.)

If $\mathcal{C}$ is structured over $\mathcal{B}$, then $A \mapsto |A|$ is the object-map of a functor $U: \mathcal{C} \to \mathcal{B}$, whose action on morphisms sends an instance $(f: A \to B)$ of the source–target predicate to f; this functor is faithful, and conversely if $U: \mathcal{C} \to \mathcal{B}$ is any faithful functor then $\mathcal{C}$ may be structured over $\mathcal{B}$ by setting $|A| = U(A)$ and

$$(f: A \to B) \quad \Leftrightarrow \quad (\exists g)(g: A \to B \text{ in } \mathcal{C} \text{ and } U(g) = f) .$$

A category structured over **Set** is called *concrete* (note that we consider concreteness as an extra structure rather than a property of a category), and a category structured over the category **1** with one object and one morphism (equivalently, a category definable by a protocategory with just one protomorphism – we usually denote the latter by $\leq$) is called a *preorder*.

A construction which we shall use a great deal in this book is that of the *slice category* $\mathcal{C}/B$, where $\mathcal{C}$ is a category and B is an object of $\mathcal{C}$. This may be defined conveniently as a structured category over $\mathcal{C}$: its objects are the morphisms of $\mathcal{C}$ with codomain B, the underlying object of $f: A \to B$ is $A = \text{dom } f$, and the source–target predicate is given by $h: f \to g$ iff $gh = f$, i.e. 'the morphisms of $\mathcal{C}/B$ are commutative triangles'. We note that in this case the forgetful functor $\mathcal{C}/B \to \mathcal{C}$ is not merely faithful but conservative: if $h: f \to g$ and h is an

isomorphism in $\mathcal{C}$ then $h^{-1}: g \to f$. We shall write $B\backslash\mathcal{C}$ for the *co-slice category* $(\mathcal{C}^{\mathrm{op}}/B)^{\mathrm{op}}$; but co-slices will occur much less frequently than slices in what follows.

Many of the categorical properties which we consider will be stable under slicing – that is, if a category $\mathcal{C}$ has the given property, then so do all its slice categories $\mathcal{C}/B$. As a general rule, given a property P of categories, we shall say that $\mathcal{C}$ is *locally* P if all its slice categories have property P. (There is a geometric intuition behind this use of the word 'locally' – see Chapter C1 – but note that the case P = 'small' is an exception to our general scheme.) We observe that, if $f: A \to B$ is an object of $\mathcal{C}/B$, then the slice category $(\mathcal{C}/B)/f$ is canonically isomorphic to $\mathcal{C}/A$; so a property of the form 'locally P' is automatically stable under slicing. Note also that if B is a terminal object of $\mathcal{C}$ then the forgetful functor $\mathcal{C}/B \to \mathcal{C}$ is an isomorphism; and that a slice category $\mathcal{C}/B$ always has a terminal object, namely the identity morphism 1_B. For this reason, we shall normally denote the terminal object of a category, if it has one, by 1; and if B is an object of $\mathcal{C}$ we shall tend to use B also as a name for the unique morphism $B \to 1$.

Example 1.1.6 If B is a set, then the slice category $\mathbf{Set}/B$ is equivalent (though not isomorphic) to the B-fold cartesian power $\mathbf{Set}^B$: in one direction the equivalence sends an object $f: A \to B$ of $\mathbf{Set}/B$ to the family of fibres $(A_b \mid b \in B)$ where $A_b = f^{-1}(b) \subseteq A$, and in the other it sends a B-indexed family $(A_b \mid b \in B)$ to the disjoint union

$$\coprod_{b \in B} A_b = \bigcup \{A_b \times \{b\} \mid b \in B\}$$

together with the obvious projection $\coprod A_b \to B$. We leave to the reader the easy verification that these constructions are functorial and inverse to each other up to natural isomorphism.

We shall often 'extend' the equivalence just described in an informal way, by thinking of objects of $\mathcal{C}/B$ (for an arbitrary $\mathcal{C}$) as 'B-indexed families of objects of $\mathcal{C}$'; the formalism underlying this idea is developed in Chapters B1 and B2.

If $\mathcal{C}$ and $\mathcal{D}$ are categories, we write $[\mathcal{C}, \mathcal{D}]$ for the *functor category* whose objects are all functors $\mathcal{C} \to \mathcal{D}$, and whose morphisms are natural transformations between them. A further example of 'stability under slicing', which in fact extends the example above, and which will be of considerable importance later on, is the following.

Proposition 1.1.7 *Let $\mathcal{C}$ be a category, and $F: \mathcal{C} \to \mathbf{Set}$ a functor. Then there is a category $\mathcal{F}$ such that the slice category $[\mathcal{C}, \mathbf{Set}]/F$ is equivalent to the functor category $[\mathcal{F}, \mathbf{Set}]$. Moreover, if $\mathcal{C}$ is small (resp. locally small), then so is $\mathcal{F}$.*

Proof $\mathcal{F}$ is most conveniently defined as a category structured over $\mathcal{C}$: its objects are the elements of the disjoint union of the sets $F(A), A \in \mathrm{ob}\,\mathcal{C}$, the

underlying object of x being the unique A such that $x \in F(A)$, and the source–target predicate is given by

$$f: x \to y \quad \text{iff} \quad F(f)(x) = y\,.$$

The fact that $\mathcal{F}$ is small (resp. locally small) if $\mathcal{C}$ is follows immediately from the definition.

Given an object $\alpha: G \to F$ of $[\mathcal{C}, \mathbf{Set}]/F$, we define a functor $\Phi(\alpha): \mathcal{F} \to \mathbf{Set}$ by

$$\Phi(\alpha)(x) = \{y \in G(A) \mid \alpha_A(y) = x\}$$

if $x \in F(A)$, with $\Phi(\alpha)(f)(y) = G(f)(y)$ whenever this makes sense. It is clear that a morphism $\gamma: (\alpha: G \to F) \to (\beta: H \to F)$ in $[\mathcal{C}, \mathbf{Set}]/F$ induces a natural transformation $\Phi(\gamma): \Phi(\alpha) \to \Phi(\beta)$, so that Φ becomes a functor $[\mathcal{C}, \mathbf{Set}]/F \to [\mathcal{F}, \mathbf{Set}]$.

In the opposite direction, given a functor $H: \mathcal{F} \to \mathbf{Set}$, let $\Psi(H)$ denote the functor $\mathcal{C} \to \mathbf{Set}$ defined by

$$\Psi(H)(A) = \coprod_{x \in FA} H(x),$$

with $\Psi(H)(f)$, for $f: A \to B$ in $\mathcal{C}$, defined as the union of the functions $H(f: x \to F(f)(x))$ over all $x \in F(A)$. $\Psi(H)$ comes equipped with an obvious natural transformation to F, and Ψ itself is a functor $[\mathcal{F}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]/F$. We leave to the reader the remaining details of the verification of this fact, and of the fact that Φ and Ψ are inverse to each other up to natural isomorphism. □

We assume familiarity with the process of 'Karoubian completion' or 'Cauchy completion', that is the splitting of idempotents. An *idempotent* in a category $\mathcal{C}$ is of course a morphism $e: A \to A$ such that $ee = e$; e is said to *split* if it can be factored as fg where gf is an identity morphism. (If such a splitting exists, then f is necessarily an equalizer of e and 1_A, and g is a coequalizer of the same pair; conversely, if $\mathcal{C}$ has either equalizers or coequalizers then all its idempotents have splittings. For the same reason, the splitting of an idempotent, if it exists, is unique up to canonical isomorphism.) If $\mathcal{C}$ is an arbitrary category and $\mathcal{E}$ is a class of idempotents of $\mathcal{C}$ (which we assume contains all identity morphisms of $\mathcal{C}$), we define $\mathcal{C}[\check{\mathcal{E}}]$ to be the category generated by the protocategory whose objects are the members of $\mathcal{E}$, whose protomorphisms are the morphisms of $\mathcal{C}$, and whose source–target predicate is given by

$$f: e \to e' \quad \text{iff} \quad (e'fe \text{ is defined and})\ e'fe = f\,.$$

The composition predicate of $\mathcal{C}[\check{\mathcal{E}}]$ is that of $\mathcal{C}$; however, $\mathcal{C}[\check{\mathcal{E}}]$ is *not* a category structured over $\mathcal{C}$ as we defined it in 1.1.5, since the identity morphism on an object e is e itself, not $1_{\mathrm{dom}\ e}$.

Lemma 1.1.8 *There is a full and faithful functor $I: \mathcal{C} \to \mathcal{C}[\check{\mathcal{E}}]$, which is universal amongst functors defined on $\mathcal{C}$ and mapping all elements of $\mathcal{E}$ to split idempotents.*

Proof We define $I(A) = 1_A$ and $I(f) = f: 1_{\mathrm{dom}\ f} \to 1_{\mathrm{cod}\ f}$; this is clearly a functor, and is full and faithful. If $e: A \to A$ is a member of $\mathcal{E}$, then

$$1_A \xrightarrow{\quad e \quad} e \xrightarrow{\quad e \quad} 1_A$$

is a splitting of I(e) in $\mathcal{E}[\check{\mathcal{E}}]$; and if $T: \mathcal{C} \to \mathcal{D}$ is any functor sending the members of $\mathcal{E}$ to split idempotents, we may (modulo some use of choice!) extend it, uniquely up to natural isomorphism, to a functor $\check{T}: \mathcal{C}[\check{\mathcal{E}}] \to \mathcal{D}$, by defining $\check{T}$(e) to be the image of a splitting of $T(e)$. □

Corollary 1.1.9 *For any category $\mathcal{C}$, the functor categories $[\mathcal{C}, \mathbf{Set}]$ and $[\mathcal{C}[\check{\mathcal{E}}], \mathbf{Set}]$ are equivalent, where $\mathcal{E}$ is the class of all idempotents of $\mathcal{C}$.*

Proof Since **Set** has (canonical) equalizers, all its idempotents split; so every functor $\mathcal{C} \to \mathbf{Set}$ extends canonically to a functor defined on $\mathcal{C}[\check{\mathcal{E}}]$. It is straightforward to verify that the extension also works for arbitrary natural transformations between such functors. □

Thus, when considering functor categories of the form $[\mathcal{C}, \mathbf{Set}]$ (as we shall do frequently throughout this book), there is no loss of generality in assuming that $\mathcal{C}$ is a category in which all idempotents split. Such categories $\mathcal{C}$ are commonly called *Cauchy-complete* or *Karoubian.* In fact, if $\mathcal{C}$ is Cauchy-complete, it can be recovered up to equivalence from $[\mathcal{C}, \mathbf{Set}]$: recall that it follows from the Yoneda lemma that the representable functors $\mathcal{C}(A, -)$ are projective in $[\mathcal{C}, \mathbf{Set}]$, and they are also *indecomposable* in the sense that they have no nontrivial coproduct (= disjoint union) decompositions.

Lemma 1.1.10 *If $\mathcal{C}$ is a small Cauchy-complete category, then $\mathcal{C}^{\mathrm{op}}$ is (weakly) equivalent to the full subcategory of $[\mathcal{C}, \mathbf{Set}]$ whose objects are indecomposable projectives.*

Proof The Yoneda embedding yields a weak equivalence from $\mathcal{C}^{\mathrm{op}}$ to the full subcategory of representable functors in $[\mathcal{C}, \mathbf{Set}]$, and we have just observed that the latter are all indecomposable projectives. Conversely, let F be an indecomposable projective. The family of all natural transformations from representable functors to F is jointly epic, and so yields an epimorphism $e: G \twoheadrightarrow F$ from a coproduct of representable functors to F. Since F is projective, this epimorphism is split, say by $d: F \rightarrowtail G$. Pulling back the coproduct decomposition of G along d yields a coproduct decomposition of F, which must be trivial; so F is a retract of a single one of the representable functors whose coproduct is G, i.e. it is the image of an idempotent endomorphism of some $\mathcal{C}(A, -)$. But since the Yoneda embedding is full and faithful, this idempotent derives from a (unique) idempotent endomorphism of A. Since $\mathcal{C}$ is Cauchy-complete, this endomorphism

splits (with image B, say); and since any functor preserves images of idempotents we have $F \cong \mathcal{C}(B, -)$. Thus the indecomposable projectives in $[\mathcal{C}, \mathbf{Set}]$ are (up to isomorphism) exactly the representable functors. □

A1.2 Cartesian categories

In much of this book we shall be concerned with particular classes of categories having certain 'set-like' categorical properties. (Note, however, that by 'set-like' we do not mean 'set-theoretic'; we are here talking about elementary properties, ones which may be formulated in the first-order language of categories.) The simplest of these, and the one on which (almost) all the others will depend, is that of having finite limits; so this section is devoted to recalling some facts about it.

Of course, having finite limits is an elementary property:

Lemma 1.2.1 *For a category $\mathcal{C}$, the following are equivalent:*

(i) *Every functor $\mathcal{D} \to \mathcal{C}$, where $\mathcal{D}$ is a finite category, has a limit.*

(ii) *$\mathcal{C}$ has finite products and equalizers of pairs of morphisms.*

(iii) *$\mathcal{C}$ has a terminal object, products of pairs of objects and equalizers of pairs of morphisms.*

(iv) *$\mathcal{C}$ has a terminal object and pullbacks of pairs of morphisms.*

Proof See any good textbook on category theory. □

Lemma 1.2.1 may be read in two different ways: as establishing the equivalence of statements of the form 'For all diagrams of type... in $\mathcal{C}$, there exists a limit', or (more constructively) as saying that if we are given an operation assigning a particular choice of limit to each diagram of one of certain specified kinds, then we can use it to construct an operation which does the same for diagrams of certain other kinds. In keeping with our desire to demand as little as possible of our metatheory, we shall generally take the latter view: when we say that $\mathcal{C}$ *has limits* of a particular kind, we mean that we are given an operation assigning limits to all diagrams of this kind, unless we explicitly indicate otherwise. This means that the existence of finite limits is, strictly speaking, an additional structure rather than a property of a category $\mathcal{C}$; but it does little harm to think of it as a property, because the structure – if it exists – is unique up to unique isomorphism (unlike the structure of concreteness, for example). In fact, most of the 'properties' we shall consider in the next few sections are really structures, but they all have the same degree of uniqueness about them.

We call a category *cartesian* if it has the structure indicated in (iii) or (iv) of Lemma 1.2.1. A functor $F: \mathcal{C} \to \mathcal{D}$ between cartesian categories is called *cartesian* if it preserves this structure (but only in the 'up-to-isomorphism' sense; it is clearly unreasonable to expect that we can make a canonical choice of limits

in two categories $\mathcal{C}$ and $\mathcal{D}$ in such a way that all interesting cartesian functors between them will respect these choices). The term 'left exact' is often used for functors preserving finite limits, but the doubly dead metaphor contained in this term (it is derived from exact differentials, by way of exact sequences) means that it is best avoided, except possibly in the particular context of abelian categories where the notion of exact sequence still has a meaning.

The category **Set** is cartesian: any one-element set, say $\{\emptyset\}$, gives us a choice of terminal object, the usual definition of cartesian product gives us a choice of binary products, and subsets give us a choice of equalizers. Many categories structured over **Set** (for example, **Gp** and the category **Sp** of topological spaces and continuous maps) are cartesian, and their cartesian structure may be chosen in such a way that the forgetful functor to **Set** preserves it strictly. If $\mathcal{C}$ is cartesian, then the category $[\mathcal{D}, \mathcal{C}]$ of functors $\mathcal{D} \to \mathcal{C}$ is cartesian for any $\mathcal{D}$, with structure 'defined pointwise' from that of $\mathcal{C}$ (that is, the functors 'evaluate at B': $[\mathcal{D}, \mathcal{C}] \to \mathcal{C}$ are cartesian for each $B \in \text{ob}\,\mathcal{D}$, and the structure may be chosen so that they preserve it strictly). If $\mathcal{C}$ is a locally small cartesian category, then for any $A \in \text{ob}\,\mathcal{C}$ the representable functor $\mathcal{C}(A, -) \colon \mathcal{C} \to \mathbf{Set}$ is cartesian; hence

Lemma 1.2.2 *For any small cartesian category $\mathcal{C}$, there exists a conservative cartesian functor $\mathcal{C} \to \mathbf{Set}^B$ (equivalently, $\mathcal{C} \to \mathbf{Set}/B$) for some set B.*

Proof Let B be the set of all objects of $\mathcal{C}$, and consider the functor whose Ath coordinate is $\mathcal{C}(A, -)$. □

Although conservative, the functor constructed in Lemma 1.2.2 is not (normally) full; but it can be made so by considering its codomain to be a functor category rather than a slice category of **Set**:

Lemma 1.2.3 *For any locally small cartesian category $\mathcal{C}$, there is a category $\mathcal{D}$ (which may be taken to be small if $\mathcal{C}$ is) and a full and faithful cartesian functor $\mathcal{C} \to [\mathcal{D}, \mathbf{Set}]$.*

Proof Take $\mathcal{D} = \mathcal{C}^{\text{op}}$ and use the Yoneda embedding. □

Both 1.2.2 and 1.2.3 can be used as the basis for 'metatheorems' which assert that any statement (of a prescribed type) which can be expressed in the first-order language of cartesian categories, and which is true is all slices of **Set** (or in all **Set**-valued functor categories, as the case may be), is true in all cartesian categories. We shall not explore this aspect of category theory in detail here.

The hypothesis of smallness in Lemma 1.2.2 can be relaxed: all we need is a set $B \subseteq \text{ob}\,\mathcal{C}$ such that the functors $\mathcal{C}(A, -)$, $A \in B$, are jointly conservative. Such a set will be called a *generating set* for $\mathcal{C}$; if B is a singleton $\{A\}$, we call A a *generator* for $\mathcal{C}$. We shall also wish to consider the weaker condition that the functors $\mathcal{C}(A, -)$, $A \in B$, are jointly faithful; in this case we call B a *separating set* (and, if $B = \{A\}$, we call A a *separator*) for $\mathcal{C}$.

The notions of separating and generating set can be defined in elementary terms (and without assuming local smallness of $\mathcal{C}$). First we note:

Lemma 1.2.4 *Suppose $\mathcal{C}$ has equalizers, and $F: \mathcal{C} \to \mathcal{D}$ preserves them. Then F is conservative iff it 'preserves properness of subobjects', i.e. iff, given a monomorphism f in $\mathcal{C}$, $F(f)$ invertible implies f invertible.*

Proof The condition is clearly necessary. Conversely, it implies that F is faithful, since given a parallel pair (f, g) in $\mathcal{C}$ we have $f = g$ iff the equalizer of (f, g) is an isomorphism. Now if f is any morphism of $\mathcal{C}$ such that $F(f)$ is an isomorphism, then given any parallel pair (g, h) with $fg = fh$ we have $F(g) = F(h)$, whence $g = h$; so f is monic, whence f is an isomorphism by the given condition. □

Thus, in the absence of any set-theoretic assumptions on a cartesian category $\mathcal{C}$, we may define a family $\mathcal{G}$ of objects of $\mathcal{C}$ to be generating if, given a monomorphism $f: A \to B$ in $\mathcal{C}$, the condition that every morphism $G \to B$ with $G \in \mathcal{G}$ factors through f forces f to be an isomorphism. We leave to the reader the task of formulating an elementary definition of separating family (if he has not already seen it). In a category $\mathcal{C}$ with equalizers, a generating family is always a separating family, by the argument contained in the proof of 1.2.4; the converse holds if $\mathcal{C}$ is balanced, but not in general.

Example 1.2.5 It is easy to see that the category **Sp** of topological spaces and continuous maps has a single separator, namely the 1-point space. However, it does not have any generating set of objects. To see this, note that for any infinite cardinal κ we can find a set X and two topologies $\mathcal{O}_1, \mathcal{O}_2$ on X such that $\mathcal{O}_1 \subset \mathcal{O}_2$ but their restrictions to any subspace of X of cardinality less than κ coincide. (For example, we may take X to be the ordinal $\lambda + 1$, where λ is a limit ordinal of cofinality at least κ, $\mathcal{O}_1$ to be the usual order topology on X and $\mathcal{O}_2$ to be generated by $\mathcal{O}_1$ together with the singleton $\{\lambda\}$.) Thus, if $\mathcal{G} = \{G_i \mid i \in I\}$ is any set of objects of **Sp**, then by taking κ larger than the cardinalities of all the G_i we obtain a monomorphism (namely the identity map $(X, \mathcal{O}_2) \to (X, \mathcal{O}_1)$) which is not an isomorphism, but such that this fact cannot be 'detected' by any of the G_i.

The forgetful functor $\mathcal{C}/B \to \mathcal{C}$ is not cartesian (unless B is a terminal object of $\mathcal{C}$), but it does preserve pullbacks (and, more generally, all limits of connected diagrams) and even creates them. Since $\mathcal{C}/B$ always has a terminal object, we deduce that it is cartesian provided $\mathcal{C}$ has pullbacks; but the converse is also true, since pullbacks over B in $\mathcal{C}$ are 'the same thing as' products in $\mathcal{C}/B$. That is,

Lemma 1.2.6 *A category is locally cartesian iff it has pullbacks.* □

In particular, the property of cartesianness is stable under slicing.

Example 1.2.7 There are not many (naturally arising) examples of locally cartesian categories which fail to be cartesian, but there is one which will be

of importance later on, namely the category **LH** of topological spaces and local homeomorphisms. This has the same objects as **Sp**, but its morphisms are those continuous maps $f: X \to Y$ such that each point $x \in X$ has an open neighbourhood which is mapped homeomorphically by f onto an open neighbourhood of $f(x)$. (An equivalent characterization is that both f and the diagonal embedding $X \to X \times_Y X$ are open maps; cf. C3.1.15.) A pullback in **Sp** of a local homeomorphism is a local homeomorphism (this is easily seen from the definition, but may also be proved using the second characterization above, and the fact that open maps are stable under pullback); from this it can be shown that **LH** has pullbacks, and the inclusion functor **LH** $\to$ **Sp** preserves them (cf. C1.3.2(iv)). So by Lemma 1.2.6 each slice category **LH**$/X$ is cartesian, and the inclusion **LH**$/X \to$ **Sp**$/X$ is a cartesian functor. However, **LH** does not have a terminal object, as may be seen by the following argument: for each infinite cardinal κ, there exists a space X_κ (for example, the product of λ copies of a discrete two-point space, where $2^\lambda \geq \kappa$) in which every nonempty open set has cardinality at least κ; now any space to which X_κ maps by a local homeomorphism – in particular, a terminal object of **LH**, if it exists – must have cardinality at least κ. (In contrast to the above, **LH** does have binary products; hence it also has equalizers, but they are not preserved by the inclusion **LH** $\to$ **Sp** – to obtain the equalizer of a pair of morphisms in **LH**, one must take the interior of the subspace which is their equalizer in **Sp**. As an example of a pair of local homeomorphisms whose equalizer in **Sp** is not an open subspace, we may take the identity mapping on $\mathbb{R}$ and the map $x \mapsto -x$.)

Let $f: A \to B$ be a morphism in an arbitrary category $\mathcal{C}$. We have seen that $\mathcal{C}/A$ is isomorphic to a slice category of $\mathcal{C}/B$, and so we have a forgetful functor $\mathcal{C}/A \to \mathcal{C}/B$, whose effect on objects is simply the map $g \mapsto fg$; we shall denote this functor by Σ_f.

Lemma 1.2.8 *$\mathcal{C}$ is locally cartesian iff Σ_f has a right adjoint f^*, for every morphism f of $\mathcal{C}$.*

Proof If $\mathcal{C}$ has pullbacks then we define $f^*(g)$, for an object g of $\mathcal{C}/B$, to be the left vertical map in the pullback square

$$\begin{tikzcd} P \arrow[r] \arrow[d] & C \arrow[d, "g"] \\ A \arrow[r, "f"] & B \end{tikzcd} \quad .$$

It is straightforward to verify that this defines a functor $\mathcal{C}/B \to \mathcal{C}/A$, and that it is right adjoint to Σ_f. Conversely, suppose f^* exists for every f. Then, given a pair of morphisms (f, g) with common codomain, the counit of the adjunction

$(\Sigma_f \dashv f^*)$ yields a morphism from the domain of $f^*(g)$ to the domain of g making the square

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d, "f^*(g)"'] & \bullet \arrow[d, "g"] \\ \bullet \arrow[r, "f"] & \bullet \end{tikzcd}$$

commute, and the universal property of the adjoint says that this square is a pullback. □

We note that the assignment $f \mapsto \Sigma_f$ is itself functorial; that is, $\Sigma_f \Sigma_g = \Sigma_{fg}$ for any composable pair (f, g). In general, the assignment $f \mapsto f^*$ is not functorial in this strict sense; we cannot expect our canonical choice of pullbacks in $\mathcal{C}$ to be such that the functors $(fg)^*$ and $g^* f^*$ are exactly equal. However, they are canonically naturally isomorphic (being right adjoints to the same functor). We shall explore this point further in Chapter B1.

Just as all finite limits can be constructed from pullbacks and a terminal object, we can construct more complicated finite limits from pullbacks alone. The class of finite limits which can be constructed from pullbacks has been characterized by R. Paré [931]: they are the limits over finite simply connected categories. (A category $\mathcal{C}$ is said to be *simply connected* if the groupoid reflection of $\mathcal{C}$ – i.e. the category obtained by freely adjoining inverses for all the morphisms of $\mathcal{C}$ – is a trivial connected groupoid; equivalently, $\mathcal{C}$ is simply connected if every functor from $\mathcal{C}$ to a groupoid is naturally isomorphic to a constant functor.) Clearly, this class does not include all limits over finite connected categories – the 'parallel-pair' category which gives rise to equalizers is connected but not simply connected – but in fact we may construct all finite connected limits from pullbacks and equalizers. However, when we turn to the preservation of limits by functors, there is a strong tendency for those which preserve pullbacks to preserve equalizers as well: the following is a well-known folk-theorem, but does not often appear in print, so we feel it is worth stating explicitly.

Lemma 1.2.9 *Let $\mathcal{C}$ be a cartesian category, and $T: \mathcal{C} \to \mathcal{D}$ a functor. If T preserves pullbacks, then it preserves all finite connected limits.*

Proof We may factor T as

$$\mathcal{C} \xrightarrow{\hat{T}} \mathcal{D}/T1 \xrightarrow{\Sigma_{T1}} \mathcal{D}$$

where 1 is the terminal object of $\mathcal{C}$, and $\hat{T}(A)$ is the image under T of the unique morphism $A \to 1$. Since Σ_{T1} creates pullbacks, $\hat{T}$ preserves them; but it also preserves the terminal object by construction, and hence preserves all finite limits. But Σ_{T1} preserves all connected limits which exist in $\mathcal{D}/T1$, and so T preserves finite connected limits. □

Note that Lemma 1.2.9 does not require any hypotheses on the existence of limits in $\mathcal{D}$. However, the hypotheses on $\mathcal{C}$ cannot be weakened to the existence of finite connected limits: as we have already seen, the inclusion functor $\mathbf{LH} \to \mathbf{Sp}$ provides an example of a functor which preserves pullbacks but not equalizers, even though both exist in $\mathbf{LH}$.

Although, as already mentioned, not all equalizers are simply connected limits, there is an important class of them which are, namely coreflexive equalizers. We have already met reflexive coequalizers in Theorem 1.1.2; dually, we say a parallel pair $f, g \colon A \rightrightarrows B$ is *coreflexive* if there exists $d \colon B \to A$ such that $df = dg = 1_A$. The equalizer of a coreflexive pair may also be regarded as the limit of a diagram of shape $\mathcal{E}$, where $\mathcal{E}$ is the finite category with two objects A and B and five non-identity morphisms $f, g \colon A \rightrightarrows B$, $d \colon B \to A$ and $u, v \colon B \rightrightarrows B$ satisfying the equations $df = dg = 1_A$, $fd = u$ and $gd = v$ (other composites being deducible from these). It is easily seen that $\mathcal{E}$ is a simply connected category, since it has a terminal object A; hence, by Paré's theorem quoted above, coreflexive equalizers may be constructed from pullbacks alone. But in fact we can be more explicit:

Lemma 1.2.10 *Let $f, g \colon A \rightrightarrows B$ be a coreflexive pair in a category $\mathcal{C}$. Then a morphism $e \colon E \to A$ is an equalizer of f and g iff the square*

$$\begin{tikzcd} E \arrow[r, "e"] \arrow[d, "e"'] & A \arrow[d, "g"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

is a pullback.

Proof Given a parallel pair $h, k \colon C \rightrightarrows A$, the equation $fh = gk$ implies $h = dfh = dgk = k$, where d is a common splitting for f and g; so a pullback of f against g has the same universal property as an equalizer of f and g. □

Coreflexive equalizers plus finite products suffice to construct all finite limits: this may be proved by observing that the standard construction of an arbitrary finite limit as an equalizer of a pair of morphisms between products in fact requires the equalizer of a coreflexive pair. But we may also note that, if $\mathcal{C}$ has finite products, then the equalizer of $f, g \colon A \rightrightarrows B$ may also be obtained as the equalizer of $(1_A, f), (1_A, g) \colon A \rightrightarrows A \times B$; and the latter pair is coreflexive, with common splitting given by the first projection.

Another important property of coreflexive equalizers is that, because the diagonal functor $\mathcal{E} \to \mathcal{E} \times \mathcal{E}$ is initial (cf. B2.5.12), we may reduce the

computation of limits of shape $\mathcal{E} \times \mathcal{E}$ to that of limits of shape $\mathcal{E}$. Explicitly, this means the following:

Lemma 1.2.11 *Suppose given a diagram*

$$\begin{tikzcd}
A_1 \arrow[r] \arrow[d] & B_1 \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & C_1 \arrow[d] \\
A_2 \arrow[r] \arrow[d, shift left] \arrow[d, shift right] & B_2 \arrow[r, shift left] \arrow[r, shift right] \arrow[d, shift left] \arrow[d, shift right] & C_2 \arrow[d, shift left] \arrow[d, shift right] \\
A_3 \arrow[r] & B_3 \arrow[r, shift left] \arrow[r, shift right] & C_3
\end{tikzcd}$$

which commutes in the appropriate sense (e.g. each of the four composites $B_2 \to C_2 \to C_3$ is equal to the corresponding composite $B_2 \to B_3 \to C_3$), and in which the rows and columns are equalizers and the parallel pairs are all coreflexive. Then the diagonal $A_1 \to B_2 \rightrightarrows C_3$ is an equalizer.

Proof It is clear that the two composites along the diagonal are equal; so suppose we are given a morphism $D \to B_2$ equalizing $B_2 \rightrightarrows C_3$. By composing with a common splitting for $C_2 \rightrightarrows C_3$, we see that $D \to B_2$ also equalizes $B_2 \rightrightarrows C_2$, and hence factors uniquely through $A_2 \to B_2$. Similarly, $D \to B_2$ equalizes $B_2 \rightrightarrows B_3$; so its factorization through A_2 equalizes $A_2 \rightrightarrows A_3$, since $A_3 \to B_3$ is monic, and hence in turn factors uniquely through $A_1 \to A_2$. So we have shown that $A_1 \to B_2$ has the universal property of an equalizer. □

Corollary 1.2.12 *Suppose given a bifunctor $\otimes : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ such that, for each $A \in \text{ob}\, \mathbf{C}$, the functors $A \otimes (-)$ and $(-) \otimes A$ preserve coreflexive equalizers. Then, if $A_i \to B_i \rightrightarrows C_i$ $(i = 1, 2)$ are coreflexive equalizer diagrams in $\mathbf{C}$, so is the diagram $A_1 \otimes A_2 \to B_1 \otimes B_2 \rightrightarrows C_1 \otimes C_2$.*

Proof Apply 1.2.11 to the diagram

$$\begin{tikzcd}
A_1 \otimes A_2 \arrow[r] \arrow[d] & A_1 \otimes B_2 \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & A_1 \otimes C_2 \arrow[d] \\
B_1 \otimes A_2 \arrow[r] \arrow[d, shift left] \arrow[d, shift right] & B_1 \otimes B_2 \arrow[r, shift left] \arrow[r, shift right] \arrow[d, shift left] \arrow[d, shift right] & B_1 \otimes C_2 \arrow[d, shift left] \arrow[d, shift right] \\
C_1 \otimes A_2 \arrow[r] & C_1 \otimes B_2 \arrow[r, shift left] \arrow[r, shift right] & C_1 \otimes C_2
\end{tikzcd} .$$

□

In practice, we shall most commonly need the duals of 1.2.11 and 1.2.12, referring to reflexive coequalizers rather than coreflexive equalizers. Note in particular that 1.2.12 contains the result that reflexive coequalizers commute with finite products in **Set** (or, more generally, in any cartesian closed category; cf. Section A1.5 below).

Another class of colimits which are well known to commute with finite products (and indeed all finite limits) in **Set** is the class of filtered colimits. We shall prove an 'internal' version of this commutation result in Section B2.6; but we note here that the class of all colimits which commute with finite products in **Set** has been characterized by C. Lair [660] as the class of *sifted* colimits. Moreover, all colimits in this class can be constructed from filtered colimits and reflexive coequalizers; see [26].

A1.3 Regular categories

Like many writers on category theory, we shall be deliberately ambiguous in our use of the word 'subobject': according to the context, a *subobject* of an object A in a category $\mathcal{C}$ will mean either a particular monomorphism of $\mathcal{C}$ with codomain A, or an isomorphism class of such monomorphisms (the notion of isomorphism being that of the category $\mathcal{C}/A$). For the present, let us adopt the former view, and write $\mathrm{Sub}(A)$ (or, if necessary, $\mathrm{Sub}_{\mathcal{C}}(A)$) for the full subcategory of $\mathcal{C}/A$ whose objects are the subobjects of A. This category is of course a preorder; we follow the usual custom of denoting its unique protomorphism by $\leq$. If $\mathcal{C}/A$ is cartesian, so is $\mathrm{Sub}(A)$ (and the inclusion $\mathrm{Sub}(A) \to \mathcal{C}/A$ is a cartesian functor); again, we shall follow tradition by calling products in $\mathrm{Sub}(A)$ *intersections*, and denoting them by $\cap$ rather than $\times$. (We shall also fall into the common abuse of denoting a subobject of A by the name of its domain, A' say, leaving the reader's imagination to supply a name for the actual monomorphism $A' \rightarrowtail A$.)

A pullback of a monomorphism is a monomorphism; so if $\mathcal{C}$ is (locally) cartesian the functors $f^* \colon \mathcal{C}/B \to \mathcal{C}/A$, defined in Lemma 1.2.8, restrict to functors $\mathrm{Sub}(B) \to \mathrm{Sub}(A)$, which we shall again denote by f^* (or occasionally by f^{-1}). However, the functor Σ_f does not similarly restrict, unless f itself happens to be a monomorphism. We shall say that $\mathcal{C}$ *has images* if we are given an operation assigning to each morphism f of $\mathcal{C}$ a subobject $\mathrm{im}\, f$ of its codomain, which is a least (in the sense of the preorder $\leq$) subobject of $\mathrm{cod}\, f$ through which f factors.

Lemma 1.3.1 *For a category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ has images.*

(ii) *For each object A of $\mathcal{C}$, the inclusion* $\mathrm{Sub}(A) \to \mathcal{C}/A$ *has a left adjoint* im.

(iii) *(if $\mathcal{C}$ has pullbacks) For each morphism $f \colon A \to B$ of $\mathcal{C}$, the pullback functor $f^* \colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ has a left adjoint $\exists_f$.*

Proof The equivalence of (i) and (ii) is immediate from the definition. For (ii) $\Rightarrow$ (iii), we define $\exists_f$ to be the composite

$$\mathrm{Sub}(A) \longrightarrow \mathcal{C}/A \xrightarrow{\Sigma_f} \mathcal{C}/B \xrightarrow{\mathrm{im}} \mathrm{Sub}(B)$$

and verify that it is left adjoint to f^*. For (iii) $\Rightarrow$ (ii), we note that applying $\exists_f$ (if it exists) to the terminal object 1_A of $\mathrm{Sub}(A)$ yields an image of f. □

The canonical morphism $\mathrm{dom}\, f \to \mathrm{dom}(\mathrm{im}\, f)$ which is the unit of the adjunction of 1.3.1(ii) has the property that it cannot factor through any proper subobject of its codomain. We call such a morphism a *cover*; more generally, a family of morphisms with common codomain (A, say) is called a *covering family* if the only subobjects of A through which every morphism in the family factors are isomorphisms. (We may use this notion to rephrase our definition of a generating family, given in the last section: a family $\mathcal{G}$ of objects is generating iff, for each object B, the family of all morphisms with codomain B and domain in $\mathcal{G}$ is covering.) In diagrams and elsewhere, we shall use the notation $f \colon A \twoheadrightarrow B$ to denote the fact that f is a cover.

Lemma 1.3.2 *Let $\mathcal{C}$ be a category with pullbacks, $f \colon A \to B$ a morphism in $\mathcal{C}$. The following are equivalent:*

(i) *f is a cover.*

(ii) *f is orthogonal to the class of monomorphisms in $\mathcal{C}$; i.e., given a commutative diagram*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "h"] \\ C \arrow[r, tail, "m"] & D \end{tikzcd}$$

where m is monic, there is a unique $k \colon B \to C$ with $kf = g$ and $mk = h$.

(iii) *(if covers in $\mathcal{C}$ are stable under pullback) The functor $f^* \colon \mathcal{C}/B \to \mathcal{C}/A$ is conservative.*

Proof (i) $\Leftrightarrow$ (ii): Condition (i) is a special case of (ii), in which the morphism h in the diagram is required to be the identity morphism on B. Conversely, the existence of the commutative square says that f factors through the subobject $h^*(m)$ of B; so if f is a cover this subobject must be an isomorphism, which means that h factors (uniquely, since m is monic) through m. Writing $k \colon B \to C$ for this factorization, we have $mkf = hf = mg$, whence $kf = g$ since m is monic.

(i) $\Leftrightarrow$ (iii): If m is any subobject of B through which f factors, then $f^*(m) \cong 1_A \cong f^*(1_B)$, so if f^* reflects isomorphisms then m must be an isomorphism.

Conversely, suppose f is a cover; let $g\colon C \to B$ be an object of $\mathcal{C}/B$, and let

$$\begin{tikzcd} P \arrow[r, "h"] \arrow[d, "k"'] & C \arrow[d, "g"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

be a pullback. Since h is a cover, the pullback along it of any proper subobject of C is a proper subobject of P; but since the forgetful functor $\mathcal{C}/B \to \mathcal{C}$ preserves and reflects monomorphisms, this says that f^* maps proper subobjects of g to proper subobjects of k. Since f^* also preserves equalizers (having a left adjoint), it is conservative by 1.2.4. $\square$

Morphisms with the property described in 1.3.2(ii) are sometimes called *strong epimorphisms* or *extremal epimorphisms*; but we shall not need any special name for them, since we shall only consider this property in cartesian categories. Note that, in a category with equalizers, a cover $f\colon A \twoheadrightarrow B$ is necessarily an epimorphism, since it cannot factor through the equalizer of any distinct pair of morphisms with domain B. In the opposite direction, a *regular epimorphism* (i.e. one which occurs as a coequalizer) is a cover, because any factorization of it through a subobject of its codomain can be shown to be a coequalizer of the same pair of morphisms.

The mere existence of images in a category is not sufficient for many of the purposes for which we want to use them; we need to know that they are stable under pullback, i.e. that if f and g are two morphisms with common codomain then $g^*(\mathrm{im}\ f) \cong \mathrm{im}(g^*(f))$ in $\mathrm{Sub}(\mathrm{dom}\ g)$. Since monomorphisms are always stable under pullback, this is clearly equivalent to saying that covers are stable under pullback, as in 1.3.2(iii). We say a category $\mathcal{C}$ is *regular* if it is cartesian and has images, and covers in $\mathcal{C}$ are stable under pullback. A functor between such categories is called *regular* if it preserves finite limits and covers (and hence also images, since a cartesian functor preserves monomorphisms).

The following property of images in a regular category will be needed in Section D1.3.

Lemma 1.3.3 *Let $f\colon A \to B$ be a morphism in a regular category. Then, for any subobjects $A' \rightarrowtail A$ and $B' \rightarrowtail B$, we have $\exists_f(A' \cap f^*(B')) \cong \exists_f(A') \cap B'$ in* $\mathrm{Sub}(B)$.

Proof We may form the diagram

$$\begin{tikzcd}
A' \cap f^*(B') \arrow[rr, two heads] \arrow[dd, tail] \arrow[dr] & & \exists_f(A') \cap B' \arrow[dd] \arrow[dr] & \\
& f^*(B') \arrow[rr] \arrow[dd, tail] & & B' \arrow[dd, tail] \\
A' \arrow[rr, -open triangle] \arrow[dr] & & \exists_f(A') \arrow[dr] & \\
& A \arrow[rr, two heads, "f"] & & B
\end{tikzcd}$$

in which the front, left and right faces are pullbacks; so the back face is also a pullback, and hence its top edge is a cover. But the diagonal of the right face is a monomorphism, so $\exists_f(A') \cap B'$ is the image of the composite morphism from top left to bottom right. □

The isomorphism in the statement of Lemma 1.3.3 is sometimes called *Frobenius reciprocity*; we shall meet it again, in a more general context, in 1.5.8 below.

Regularity, like cartesianness, has good stability properties: if $\mathcal{C}$ is regular, then so are any functor category $[\mathcal{D}, \mathcal{C}]$ and any slice category $\mathcal{C}/B$. Moreover, the functors 'evaluate at B': $[\mathcal{D}, \mathcal{C}] \to \mathcal{C}$ are regular (which tells you how to prove the first assertion) and the forgetful functor $\mathcal{C}/B \to \mathcal{C}$ preserves covers (which tells you how to prove the second). A preorder is regular iff it is cartesian: every morphism in a preorder is monic (and epic), and so the only covers are isomorphisms.

The category **Set** is regular: covers are surjective functions, and images are the usual set-theoretic ones. The same remarks apply to **Gp** or, more generally, to any category monadic over **Set**. Note that in **Set** and **Gp** the covers coincide with the epimorphisms, but in other cases they do not – for example in the category **Mon** of monoids, the inclusion of the additive monoid of natural numbers in the group of integers is an epimorphism which is not a cover. The category **Sp** has images: monomorphisms are injective continuous functions, and covers are surjections $X \to Y$ such that Y is topologized as a quotient space of X (these are exactly the regular epimorphisms in **Sp**, though they are not all the epimorphisms). But **Sp** is not regular, because the class of quotient maps is not stable under pullback (cf. [276]). However, **LH** is locally regular: because

local homeomorphisms are open maps, it is easily seen that any surjective local homeomorphism is a cover (in **Sp**, and hence also in **LH**); and any local homeomorphism can be factored as a surjective local homeomorphism followed by an open inclusion.

The reader will have observed from the above examples that, whilst covers do not always coincide with epimorphisms, they have a strong tendency to coincide with regular epimorphisms. Indeed, we have

Proposition 1.3.4 *In a regular category, the covers are exactly the regular epimorphisms.*

Proof We have already noted that regular epimorphisms are always covers. Conversely, let $f\colon A \twoheadrightarrow B$ be a cover in a regular category, and let $(a,b)\colon R \rightrightarrows A$ be its kernel-pair (i.e. the pullback of f against itself). We shall show that f is a coequalizer of a and b. Let c: $A \to C$ be any morphism with $ca = cb$, and let

$$A \xrightarrow{\;d\;} D \xrightarrow{(g,h)} B \times C$$

be the image factorization of $(f,c)\colon A \to B \times C$. We shall show that g is an isomorphism, so that $hg^{-1}\colon B \to C$ is a factorization of c through f (which is clearly unique, since covers are epimorphisms). To this end, it suffices to show that g is monic, since the cover f factors through it. So suppose $(k,l)\colon E \rightrightarrows D$ are such that $gk = gl$. Form the pullback

$$\begin{tikzcd} P \arrow[r, "p"] \arrow[d, "{(m,n)}"] & E \arrow[d, "{(k,l)}"] \\ A \times A \arrow[r, "d \times d"] & D \times D \end{tikzcd}$$

Now $fm = gdm = gkp = glp = gdn = fn$, so (m,n) factors through (a,b) by a morphism $q\colon P \to R$. Thus we have

$$hkp = hdm = cm = caq = cbq = cn = hdn = hlp\,.$$

But $d \times d$ is a cover, since it is the composite of $1_A \times d\colon A \times A \to A \times D$ and $d \times 1_D\colon A \times D \to D \times D$, both of which are pullbacks of d (it is easy to verify that a composite of covers is a cover, using 1.3.2(ii)), and hence p is a cover and in particular epic. So we deduce that $hk = hl$, and hence $(g,h)k = (g,h)l\colon E \to B \times C$. But (g,h) is monic by definition, so $k = l$, which completes the proof. □

Earlier writers on regular categories (for example, [68], [419]) tended to use regular epimorphisms rather than covers in formulating the definition. The above lemma shows that their definition is equivalent to the one we have adopted:

Scholium 1.3.5 *Let $\mathcal{C}$ be a cartesian category with coequalizers. Then $\mathcal{C}$ is regular iff regular epimorphisms are stable under pullback in $\mathcal{C}$.*

Proof One direction follows directly from 1.3.4. Conversely, suppose regular epimorphisms are stable under pullback. Given an arbitrary morphism $f\colon A \to B$ in $\mathcal{C}$, form its kernel-pair $(a, b)\colon R \rightrightarrows A$ and the coequalizer $d\colon A \to D$ of a and b; then since $fa = fb$ we have an induced factorization $f = gd$. Arguing (almost) exactly as in the proof of 1.3.4, we may show that g is monic; thus we have factored f as a regular epimorphism (in particular, a cover) followed by a monomorphism, and so $\mathcal{C}$ has images. Moreover, if f itself is a cover, then the monic part g of the above factorization is an isomorphism, and so f is a coequalizer of a and b; hence covers are stable under pullback in $\mathcal{C}$, since they coincide with regular epimorphisms. □

Of course, a regular category need not have coequalizers for arbitrary pairs of morphisms; the most we can deduce from the definition is that it has coequalizers for those pairs which occur as kernel-pairs (and we note that the proof of 1.3.5 still works if we assume only that these coequalizers exist). A pair of morphisms which occurs as a kernel-pair has a number of special properties of interest:

Definition 1.3.6 Let $(a, b)\colon R \rightrightarrows A$ be a parallel pair of morphisms in a cartesian category.

(*a*) We say (a, b) is a *relation* if $(a, b)\colon R \to A \times A$ is monic.

(*b*) We say (a, b) is *reflexive* if there exists $r\colon A \to R$ with $ar = br = 1_A$ (cf. 1.2.2).

(*c*) We say (a, b) is *symmetric* if there exists $s\colon R \to R$ with $as = b$ and $bs = a$.

(*d*) We say (a, b) is *transitive* if there exists $t\colon P \to R$, where P is the pullback

$$\begin{tikzcd} P \arrow[r, "q"] \arrow[d, "p"] & R \arrow[d, "a"] \\ R \arrow[r, "b"] & A \end{tikzcd} ,$$

such that $at = ap$ and $bt = bq$.

(*e*) We say (a, b) is an *equivalence relation* if it has all four of the above properties. (Note that if (a, b) is a relation, then the morphisms r, s and t which verify the other three properties are unique if they exist.)

It is easily verified that the kernel-pair of any morphism is an equivalence relation. We say an arbitrary equivalence relation is *effective* if it occurs as a kernel-pair; in a regular category, an effective equivalence relation is necessarily the kernel-pair of its own coequalizer. By an *effective regular category*, we mean a regular category in which every equivalence relation is effective. (The old name for this is 'exact category (in the sense of Barr)', but the word 'exact' has been so badly overworked by category-theorists that it is now best avoided, except in the context of abelian categories.)

Not every regular category is effective. Consider the category **TF** of torsion-free abelian groups; it is easily verified that the inclusion functor from **TF** to the category **Ab** of all abelian groups has a left adjoint and preserves regular epimorphisms, from which one may deduce that **TF** is regular and cocartesian. But the group

$$R = \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid a \equiv b \bmod 2\},$$

together with its two projections to $\mathbb{Z}$, is an equivalence relation in **TF** (because it is one in **Ab**) which is not effective; its coequalizer in **TF** is the unique homomorphism $\mathbb{Z} \to 1$, whose kernel-pair is the whole of $\mathbb{Z} \times \mathbb{Z}$. The reason is, of course, that we have excluded the torsion group $\mathbb{Z}/2\mathbb{Z}$, which 'ought to be' the coequalizer of $R \rightrightarrows \mathbb{Z}$, from membership of the category **TF**. In Section A3.3 we shall develop a technique for 'effectivizing' a regular category, essentially by adjoining the 'missing quotients' which should correspond to non-effective equivalence relations, which if applied to the particular category **TF** will produce **Ab** (at least up to equivalence). Note that **Ab** itself is effective regular, as is any category monadic over **Set**. Also, if $\mathcal{C}$ is effective regular, so are any functor category $[\mathcal{D}, \mathcal{C}]$ and any slice category $\mathcal{C}/B$: kernel-pairs and coequalizers of equivalence relations are both constructed pointwise in $[\mathcal{D}, \mathcal{C}]$ and created by the forgetful functor $\mathcal{C}/B \to \mathcal{C}$.

Example 1.3.7 Not all equivalence relations in **LH** are effective. Let S^1 denote the unit circle, let R be the disjoint union of copies of S^1 indexed by the integers, let $a\colon R \to S^1$ be the codiagonal map (the map whose restriction to each copy of S^1 in R is the identity) and $b\colon R \to S^1$ the map which rotates the nth copy of S^1 in R by $n\theta$, where θ is an irrational multiple of π. Then it is easily verified that (a and b are local homeomorphisms and) (a,b) is an equivalence relation, but there is no local homeomorphism $f\colon S^1 \to X$ even satisfying $fa = fb$, let alone having (a,b) as its kernel-pair. However, this is all that can go wrong: if an equivalence relation (a,b) in **LH** is coequalized by some local homeomorphism, it is straightforward to verify that the coequalizer of (a,b) in **Sp** lies in **LH**, and that (a,b) is its kernel-pair (in **Sp**, and hence in **LH**). So we deduce that the slice categories **LH**$/X$ are all effective regular.

In a regular category, we shall normally consider projectivity with respect to covers; that is, we shall say an object P is *projective* if, given any diagram

$$\begin{tikzcd} & P \arrow[d, "g"] \\ A \arrow[r, two heads, "f"] & B \end{tikzcd}$$

where f is a cover, there exists $h: P \to A$ with $fh = g$. Because covers are stable under pullback, this is equivalent to saying that every cover with codomain P is split epic; if $\mathcal{C}$ is locally small, it is also equivalent to saying that $\mathcal{C}(P, -): \mathcal{C} \to \mathbf{Set}$ is a regular functor.

The *support* of an object A in a regular category is the image of the unique morphism $A \to 1$. We say A is *well-supported* if its support is 1, i.e. if $A \to 1$ is a cover. And we say a regular category $\mathcal{C}$ is *capital* if 1 is a generator for the full subcategory of well-supported objects of $\mathcal{C}$.

Lemma 1.3.8 *In a capital regular category, the terminal object is projective.*

Proof We have to show that every cover $A \twoheadrightarrow 1$ is split epic. If $A \cong 1$ there is nothing to prove; if not, then the two projections $A \times A \rightrightarrows A$ are not equal and so (since $A \times A$ is well-supported if A is) there exists a morphism $1 \to A \times A$ not factoring through their equalizer $A \rightarrowtail A \times A$. Composing this with one of the projections yields a morphism $1 \to A$. □

In Section D1.5 we shall develop a technique for embedding an arbitrary small regular category (conservatively) in a capital one; using this, and the Lemma just proved, we shall be able to show that every such category admits a conservative regular functor to $\mathbf{Set}/B$ for some B. (In fact this latter result has a strengthening, just as 1.2.2 was strengthened by 1.2.3: every small regular category admits a full and faithful regular functor to a functor category $[\mathcal{D}, \mathbf{Set}]$ where $\mathcal{D}$ is small. But we shall not prove the stronger result in this book: it may be found in [68] or [145] – the latter gives a proof very similar in spirit to the arguments which we shall deploy in Section D1.5.)

We conclude this section by describing a construction for embedding an arbitrary cartesian category $\mathcal{C}$ in a regular category $\mathbf{Reg}(\mathcal{C})$.

The basic ideas underlying the construction are

(i) there should be a full and faithful functor $I: \mathcal{C} \to \mathbf{Reg}(\mathcal{C})$, preserving the cartesian structure of $\mathcal{C}$ (though not the regular structure, if $\mathcal{C}$ happens to have it already);

(ii) each object of $\mathbf{Reg}(\mathcal{C})$ should be the image of a morphism of the form $I(f)$;

(iii) each object of the form $I(A)$ should be projective in $\mathbf{Reg}(\mathcal{C})$.

(We shall give some justification for the third requirement in 1.3.10(a) below.) These ideas lead us to the following definition: the objects of $\mathbf{Reg}(\mathcal{C})$ are the morphisms of $\mathcal{C}$, and morphisms $(f_1 \colon A_1 \to B_1) \to (f_2 \colon A_2 \to B_2)$ in $\mathbf{Reg}(\mathcal{C})$ are equivalence classes of morphisms $g \colon A_1 \to A_2$ in $\mathcal{C}$ such that $f_2 g$ coequalizes the kernel-pair of f_1, two such morphisms g_1 and g_2 being equivalent if they have equal composites with f_2 (equivalently, factor through its kernel-pair). We write $[g]$ for the equivalence class of g. It is straightforward to verify that this equivalence relation is respected by composition in $\mathcal{C}$, and hence $\mathbf{Reg}(\mathcal{C})$ is a category. (We could equivalently think of the morphisms $f_1 \to f_2$ in $\mathbf{Reg}(\mathcal{C})$ as individual morphisms $A_1 \to B_2$ in $\mathcal{C}$ which coequalize the kernel-pair of f_1 and admit at least one factorization through f_2; but the description given first is more convenient for defining composition.) The functor $I \colon \mathcal{C} \to \mathbf{Reg}(\mathcal{C})$ sends an object A to the identity morphism 1_A; it is again straightforward to verify that this is full and faithful. It is also easy to see that $I(1_{\mathcal{C}})$ is a terminal object of $\mathbf{Reg}(\mathcal{C})$.

To form the pullback of a pair of morphisms $[g_1] \colon f_1 \to f_3$, $[g_2] \colon f_2 \to f_3$ in $\mathbf{Reg}(\mathcal{C})$ (where $f_i \colon A_i \to B_i$ in $\mathcal{C}$), we form the pullback

$$\begin{tikzcd} P \arrow[r, "p_1"] \arrow[d, "p_2"'] & A_1 \arrow[d, "f_3 g_1"] \\ A_2 \arrow[r, "f_3 g_2"] & B_3 \end{tikzcd}$$

in $\mathcal{C}$, and define f_4 to be $(f_1 p_1, f_2 p_2) \colon P \to B_1 \times B_2$. (Note that this construction depends only on the equivalence classes $[g_1]$ and $[g_2]$, and not on the particular representatives g_1 and g_2.) It is clear that p_1 and p_2 represent morphisms $f_4 \to f_1$ and $f_4 \to f_2$ in $\mathbf{Reg}(\mathcal{C})$, and that the square

$$\begin{tikzcd} f_4 \arrow[r, "{[p_1]}"] \arrow[d, "{[p_2]}"'] & f_1 \arrow[d, "{[g_1]}"] \\ f_2 \arrow[r, "{[g_2]}"] & f_3 \end{tikzcd}$$

commutes; to show that the square has the universal property of a pullback, suppose given $[h_1] \colon f_5 \to f_1$ and $[h_2] \colon f_5 \to f_2$ with $[g_1 h_1] = [g_2 h_2]$. Then $f_3 g_1 h_1 = f_3 g_2 h_2$, so the pair (h_1, h_2) factors through the pullback P (say by $h_3 \colon A_5 \to P$); and h_3 defines a morphism $[h_3] \colon f_5 \to f_4$ making the appropriate diagram commute. Moreover, it is easy to verify that the equivalence class $[h_3]$ is independent of the choice of representatives h_1 and h_2, and that it is the unique factorization of $([h_1], [h_2])$ through $([p_1], [p_2])$ in $\mathbf{Reg}(\mathcal{C})$.

Next, we note that a morphism $[g]: f_1 \to f_2$ in $\mathbf{Reg}(\mathcal{C})$ is monic iff the kernel-pair of f_2g coincides with that of f_1. We claim that if a morphism $[e]: f_1 \to f_2$ is representable by a split epimorphism e in $\mathcal{C}$, then it is a cover in $\mathbf{Reg}(\mathcal{C})$ (this is not a necessary condition for being a cover, but we shall see that it is satisfied by the first half of the image factorization of an arbitrary morphism of $\mathbf{Reg}(\mathcal{C})$, as constructed above); and this condition is stable under pullback, for if we have a pullback square

$$\begin{tikzcd}
(f_4 \colon P \to B_1 \times B_2) \arrow[r, "{[p_1]}"] \arrow[d, "{[p_2]}"] & (f_1 \colon A_1 \to B_1) \arrow[d, "{[e]}"] \\
(f_3 \colon A_3 \to B_3) \arrow[r, "{[g]}"] & (f_2 \colon A_2 \to B_2)
\end{tikzcd}$$

and e is split (in $\mathcal{C}$) by $m: A_2 \rightarrowtail A_1$, then p_1 is split by the factorization of $(mg, 1_{A_3}): A_3 \to A_1 \times A_3$ through the pullback P. To prove the claim, suppose e is split epic in $\mathcal{C}$ and we have a diagram

$$\begin{tikzcd}
(f_1 \colon A_1 \to B_1) \arrow[r, "{[h]}"] \arrow[dr, "{[e]}"'] & (f_3 \colon A_3 \to B_3) \arrow[d, "{[g]}"] \\
 & (f_2 \colon A_2 \to B_2)
\end{tikzcd}$$

where $[g]$ is monic. Then we claim that the composite $hm: A_2 \to A_3$ (where m is any splitting for e in $\mathcal{C}$) defines a morphism $f_2 \to f_3$ in $\mathbf{Reg}(\mathcal{C})$. For if (a, b) denotes the kernel-pair of f_2, then we have

$$f_2ghma = f_2ema = f_2a = f_2b = f_2emb = f_2ghmb$$

since the commutativity of the diagram says that $f_2e = f_2gh$, and hence $f_3hma = f_3hmb$ since $[g]$ is monic, and so the kernel-pair of f_2g coincides with that of f_3. It is now straightforward to verify that $[g]$ and $[hm]$ are inverse isomorphisms in $\mathbf{Reg}(\mathcal{C})$; so $[e]$ is a cover, as claimed.

Now we define the image factorization of an arbitrary $[g]: f_1 \to f_2$ to be

$$(f_1 : A_1 \to B_1) \xrightarrow{[1_{A_1}]} (f_2g : A_1 \to B_2) \xrightarrow{[g]} (f_2 : A_2 \to B_2)$$

(again, note that this construction depends only on the equivalence class of g); it is clear from what we have said above that the first factor of this factorization is a cover, and the second factor is monic. So every morphism of $\mathbf{Reg}(\mathcal{C})$ admits a cover–monic factorization; and this factorization is stable under pullback.

We have thus proved most of

Theorem 1.3.9 *For any cartesian category $\mathcal{C}$, there exists a regular category* $\mathbf{Reg}(\mathcal{C})$ *and a full cartesian embedding* $I\colon \mathcal{C} \to \mathbf{Reg}(\mathcal{C})$ *such that all objects* $I(A)$ *are projective in* $\mathbf{Reg}(\mathcal{C})$*, and any cartesian functor* $F\colon \mathcal{C} \to \mathcal{D}$ *from* $\mathcal{C}$ *to a regular category* $\mathcal{D}$ *extends, uniquely up to canonical isomorphism, to a regular functor* $\widetilde{F}\colon \mathbf{Reg}(\mathcal{C}) \to \mathcal{D}$.

Proof We have shown that $\mathbf{Reg}(\mathcal{C})$ is a regular category and I is a full embedding; also I preserves the terminal object by construction, and whilst it does not preserve pullbacks 'on the nose' as we constructed them in $\mathbf{Reg}(\mathcal{C})$, it is easy to see that if

$$\begin{tikzcd} P \arrow[r] \arrow[d] & A_1 \arrow[d] \\ A_2 \arrow[r] & A_3 \end{tikzcd}$$

is a pullback in $\mathcal{C}$, then $I(P)$ is isomorphic to the pullback $(P \to A_1 \times A_2)$ as constructed above in $\mathbf{Reg}(\mathcal{C})$ (the isomorphism being represented by the identity morphism on P, in both directions). For the projectivity assertion, suppose given a diagram

$$\begin{tikzcd} & I(A_1) \arrow[d, "{[g]}"] \\ (f_2\colon A_2 \to B_2) \arrow[r, "{[e]}"] & (f_3\colon A_3 \to B_3) \end{tikzcd}$$

in $\mathbf{Reg}(\mathcal{C})$, where $[e]$ is a cover. On replacing $[e]$ by the first half of its cover–monic factorization if necessary, we may assume that e is actually a split epimorphism in $\mathcal{C}$, with splitting $m\colon A_3 \rightarrowtail A_2$, say. Then mg represents a morphism $I(A_1) \to f_2$ in $\mathbf{Reg}(\mathcal{C})$, and we have $[e][mg] = [g]$.

For the final assertion, we define $\widetilde{F}(f\colon A \to B)$ to be the image of $Ff\colon FA \to FB$ in $\mathcal{D}$, and if $[g]\colon f_1 \to f_2$ is a morphism of $\mathbf{Reg}(\mathcal{C})$ we define $\widetilde{F}([g])$ to be the factorization of the composite

$$FA_1 \xrightarrow{Fg} FA_2 \twoheadrightarrow \widetilde{F}(f_2)$$

through $FA_1 \twoheadrightarrow \widetilde{F}f_1$ (which exists by the condition in the definition of morphisms of $\mathbf{Reg}(\mathcal{C})$, plus 1.3.4 and the fact that F preserves kernel-pairs). The fact that this definition is independent of the choice of the representative g follows easily from the fact that $\widetilde{F}(f_2) \to FB_2$ is monic. The verification that $\widetilde{F}$ is a regular functor, and that it is (up to canonical isomorphism) the unique regular functor extending F, are easy exercises which we leave to the reader. □

Remarks 1.3.10 (*a*) If $\mathcal{C}$ is locally small, then the Yoneda embdding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ extends to a regular functor $\mathbf{Reg}(\mathcal{C}) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. It is easy to see that this extension, like the Yoneda embedding itself, is full and faithful; and its image is weakly equivalent to the full subcategory of objects which are images in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ of morphisms between representable functors. So we could take the latter category as an alternative (non-elementary) definition of $\mathbf{Reg}(\mathcal{C})$ in the case when $\mathcal{C}$ is locally small. This also explains why it was reasonable to expect the objects in the image of I to be projective, since the representables are projective in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ (cf. 1.1.10).

(*b*) It is easy to characterize the regular categories $\mathcal{C}$ which are equivalent to one of the form $\mathbf{Reg}(\mathcal{D})$ for some $\mathcal{D}$; they are exactly those regular categories which 'have enough projectives' (that is, for every object A, there exists a cover $P \twoheadrightarrow A$ with P projective), in which the full subcategory of projective objects is closed under finite limits, and in which every object admits a monomorphism into a projective. (The first condition does not imply the second; but it does imply that the full subcategory $\mathcal{P}$ of projective objects is weakly cartesian in the sense defined in (*e*) below, since for any finite diagram in $\mathcal{P}$ we may obtain a weak limit by taking a projective cover of the limit in $\mathcal{C}$.)

(c) If $\mathcal{C}$ is already regular, then 1.3.9 implies that it is a retract of $\mathbf{Reg}(\mathcal{C})$, since the identity functor on $\mathcal{C}$ extends to a unique regular functor $\mathbf{Reg}(\mathcal{C}) \to \mathcal{C}$. Of course, the latter sends an object $f\colon A \to B$ of $\mathbf{Reg}(\mathcal{C})$ to the image of f in $\mathcal{C}$. In general, this functor is very far from being an equivalence (cf. (*d*) below); but we note that it is so iff every object of $\mathcal{C}$ is projective (equivalently, every cover in $\mathcal{C}$ is split epic). For if the retraction $\mathbf{Reg}(\mathcal{C}) \to \mathcal{C}$ is an equivalence, then its inverse I must preserve covers, whence all covers in $\mathcal{C}$ are split epic; conversely, if this holds, then an arbitrary object $f\colon A \to B$ of $\mathbf{Reg}(\mathcal{C})$ is isomorphic to $I(B')$, where $B' \rightarrowtail B$ is the image of f in $\mathcal{C}$. Thus, for example, we have $\mathbf{Reg}(\mathbf{Set}_f) \simeq \mathbf{Set}_f$, and if we assume the axiom of choice (cf. Section D4.5), then $\mathbf{Reg}(\mathbf{Set}) \simeq \mathbf{Set}$.

(*d*) The apparent simplicity of the construction of $\mathbf{Reg}(\mathcal{C})$ is somewhat deceptive; it can produce a category very much 'larger' than the category $\mathcal{C}$ from which we start. As an example, we note that *for any object A of $\mathcal{C}$, the preorder* $\mathrm{Sub}_{\mathbf{Reg}(\mathcal{C})}(I(A))$ *is equivalent to the preorder reflection of the slice category* $\mathcal{C}/A$, i.e. the quotient of the latter by the equivalence relation which identifies all parallel pairs of morphisms. For a morphism $[g]\colon (f\colon B \to C) \to I(A)$ is monic iff the kernel-pair of g coincides with that of f, and it is then isomorphic in $\mathrm{Sub}(I(A))$ to $[g]\colon g \rightarrowtail I(A)$; conversely, every object of $\mathcal{C}/A$ gives rise to a subobject of $I(A)$ in this way, and we have $[g] \leq [h]$ in $\mathrm{Sub}(I(A))$ iff there exists a morphism $g \to h$ in $\mathcal{C}/A$. Now if (for example) we take $\mathcal{C}$ to be the functor category $[\mathcal{D}, \mathbf{Set}]$ where $\mathcal{D}$ is the finite category $(\bullet \rightrightarrows \bullet)$ (we may think of the objects of $\mathcal{C}$ as *directed graphs* $(A \rightrightarrows V)$, where V is the set of vertices of the graph and A the set of arrows, and the two functions assign to each arrow its source and its target), then it is not hard to see that, for any directed graph G, the preorder $\mathrm{Sub}_{\mathcal{C}}(G)$ is equivalent to a small category (in the terminology to be

introduced in the next section, $\mathcal{C}$ is well-powered). But the preorder reflection of $\mathcal{C}$ contains a proper class of non-isomorphic objects, and thus $\mathbf{Reg}(\mathcal{C})$ is not well-powered. To see this, for each ordinal α define G_α to be the graph of the strict order-relation on α: that is, the vertices of G_α are the ordinals less than α and there is one arrow $\beta \to \gamma$ iff $\beta < \gamma$. Then it is easy to see that a graph morphism $G_\alpha \to G_\beta$ corresponds to a strictly order-preserving function from α to β; in particular, such morphisms exist iff $\alpha \leq \beta$. So $\mathrm{Sub}_{\mathbf{Reg}(\mathcal{C})}(1)$ contains a copy of the ordered class **On** of all ordinals.

(*e*) It is of interest that the construction of $\mathbf{Reg}(\mathcal{C})$ works (with only minor modifications to the description above), and produces a regular category, even if we merely assume that $\mathcal{C}$ is 'weakly cartesian' (i.e., for every finite diagram in $\mathcal{C}$, there exists a 'weak limit cone' satisfying the existence clause in the definition of a limit, but not the uniqueness). We shall not pursue the matter here; the interested reader is referred to [231].

(*f*) It is also possible to combine the 'regularization' construction $\mathcal{C} \mapsto \mathbf{Reg}(\mathcal{C})$ with the effectivization construction $\mathcal{C} \mapsto \mathbf{Eff}(\mathcal{C})$ for regular categories, to be described in Section A3.3, so as to yield a 'single-step' construction of an effective regular category from a (weakly) cartesian category. Again, we shall not go into the details here, since it is sufficient to have the two separate constructions available (but see [221], and also 3.3.13(i) below). We note that there is a significant difference between the two constructions: regularization adds new images for all the morphisms of $\mathcal{C}$ (except for those which factor as a split epimorphism followed by a monomorphism), and so even if $\mathcal{C}$ is already regular it may be very different from $\mathbf{Reg}(\mathcal{C})$ (cf. the example mentioned under (*d*) above), whereas effectivization preserves the coequalizers of all those equivalence relations that are already effective, so that if $\mathcal{C}$ is effective regular then it is equivalent to $\mathbf{Eff}(\mathcal{C})$.

Suggestions for further reading: Barr [68], Borceux [145], Carboni [220], Carboni & Celia Magno [221], Carboni & Vitale [231], Freyd & Scedrov [381], Pedicchio & Rosický [944].

A1.4 Coherent categories

We have seen that the cartesianness of a category $\mathcal{C}$ implies the cartesianness of the preorders $\mathrm{Sub}(A)$, $A \in \mathrm{ob}\,\mathcal{C}$; the next level of structure we require is that they should be cocartesian as well. Of course, for a preorder to be cocartesian, all we need is the existence of finite coproducts (which, as usual, we shall call *unions* in this context). For this, a sufficient condition (for a regular category $\mathcal{C}$) is the existence of finite coproducts in $\mathcal{C}$: if $\mathcal{C}$ has colimits of a particular kind, then so does $\mathcal{C}/A$ (and the forgetful functor $\mathcal{C}/A \to \mathcal{C}$ creates them), and if $\mathcal{C}$ is regular then $\mathrm{Sub}(A)$ is reflective in $\mathcal{C}/A$ (1.3.1(ii)) and thus inherits any colimits which exist in $\mathcal{C}/A$. Explicitly, this says that we may form the union of two

subobjects $m: A' \to A$ and $m': A'' \to A$ by forming the coproduct $A' + A''$ and then taking the image of the morphism $A' + A'' \to A$ induced by m and m'.

However, as with image factorizations, the mere existence of unions in $\mathrm{Sub}(A)$ doesn't get us very far; we want them to be stable under pullback. We define a *coherent category* to be a regular category in which each $\mathrm{Sub}(A)$ has finite unions and each $f^*: \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ preserves them. (Freyd and Scedrov [381] call such a category a pre-logos, and Makkai and Reyes [790] call it a logical category; our name is based on the fact that these categories have exactly the structure needed to interpret coherent logic – cf. D1.2.6.)

Somewhat surprisingly, the assumption of coherence has implications for the existence and properties of colimits in $\mathcal{C}$ itself. We begin with the initial object: let 1 denote the terminal object of $\mathcal{C}$, and let 0 be (the domain of) the smallest subobject of 1, i.e. the initial object of $\mathrm{Sub}(1)$. (This notation will be standard from now on.)

Lemma 1.4.1 *In a coherent category $\mathcal{C}$,*

(i) *any morphism with codomain 0 is an isomorphism;*

(ii) *0 is an initial object of $\mathcal{C}$.*

Proof (i) By assumption, if A is any object of $\mathcal{C}$, we have a pullback

$$\begin{tikzcd} 0_A \arrow[r] \arrow[d, tail] & 0 \arrow[d, tail] \\ A \arrow[r] & 1 \end{tikzcd}$$

where 0_A is (the domain of) the smallest subobject of A. But if $A \to 1$ factors through $0 \rightarrowtail 1$, then the pullback of $0 \rightarrowtail 1$ along it is the whole of A; so this condition implies that A has no proper subobjects. And if A has a morphism to 0, so does $A \times A$; so the diagonal $A \rightarrowtail A \times A$ is an isomorphism, or (equivalently) the two projections $A \times A \to A$ are equal, or equivalently again $A \to 1$ is monic. But this means that A is a subobject of 1 contained in 0; so $A \to 0$ must be an isomorphism.

(ii) From (i), it follows that the projection $0 \times A \to 0$ is an isomorphism for any A, and hence there exists a morphism $0 \to 0 \times A \to A$ for any A. But any two morphisms $0 \rightrightarrows A$ must be equal, since their equalizer is an isomorphism.
□

In any category, an initial object with the extra property described in 1.4.1(i) is called *strict*; thus Lemma 1.4.1 says that any coherent category has a strict initial object. (Note, incidentally, that the deduction of (ii) from (i) in 1.4.1 used only the cartesianness of $\mathcal{C}$.)

Next, we look at binary unions. We shall frequently require

Lemma 1.4.2 *Let A_1, A_2, A_3 be three subobjects of an object A in a coherent category. The distributive law*

$$A_1 \cap (A_2 \cup A_3) \cong (A_1 \cap A_2) \cup (A_1 \cap A_3)$$

holds in $\mathrm{Sub}(A)$.

Proof The operation $A_1 \cap (-)$ may be considered as the composite

$$\mathrm{Sub}(A) \xrightarrow{m^*} \mathrm{Sub}(A_1) \xrightarrow{\exists_m} \mathrm{Sub}(A)$$

where m is the name of the monomorphism $A_1 \rightarrowtail A$ (recall that $\exists_m$ is just the restriction of Σ_m if m is monic). But m^* preserves binary unions by assumption, and $\exists_m$ does so because it is left adjoint to m^*. □

The proof of the next result may appear somewhat complicated at first. It can be (at least formally) simplified by recasting in terms of categorical logic, once we know that coherent categories provide a sound interpretation of coherent logic (see D1.3.14).

Proposition 1.4.3 *Let $m_1 \colon A_1 \rightarrowtail A$ and $m_2 \colon A_2 \rightarrowtail A$ be subobjects in a coherent category $\mathcal{C}$. Then the square*

$$\begin{tikzcd}
A_1 \cap A_2 \arrow[r, tail, "n_1"] \arrow[d, tail, "n_2"] & A_1 \arrow[d, tail, "p_1"] \\
A_2 \arrow[r, tail, "p_2"] & A_1 \cup A_2
\end{tikzcd}$$

is both a pullback and a pushout in $\mathcal{C}$.

Proof That the square is a pullback is true in any cartesian category with unions, since it is obtained by applying the pullback-preserving functor $\mathrm{Sub}(A) \to \mathcal{C}/A \to \mathcal{C}$ to a pullback in $\mathrm{Sub}(A)$. The interest of the proposition thus lies in the fact that it is a pushout, since the inclusion $\mathrm{Sub}(A) \to \mathcal{C}/A$ does not generally preserve colimits.

To prove the result, let us simplify our notation by writing B, C for $A_1 \cap A_2$, $A_1 \cup A_2$ respectively, and suppose we are given a pair of morphisms $f_1 \colon A_1 \to D$, $f_2 \colon A_2 \to D$ with $f_1 n_1 = f_2 n_2$. The morphisms $(p_1, f_1) \colon A_1 \to C \times D$ and $(p_2, f_2) \colon A_2 \to C \times D$ are both monic, since p_1 and p_2 are; let $(g, h) \colon E \to C \times D$ be their union in $\mathrm{Sub}(C \times D)$. We aim to show that g is an isomorphism, so that $hg^{-1} \colon C \to D$ is a factorization of (f_1, f_2) through (p_1, p_2) – if so, it is clearly unique, since the equalizer of two distinct factorizations would be a proper subobject of C containing both A_1 and A_2.

Let $\pi: C \times D \to C$ be the projection. Since $\exists_\pi : \mathrm{Sub}(C \times D) \to \mathrm{Sub}(C)$ is left adjoint to π^*, it preserves unions, so

$$\exists_\pi(g,h) = \exists_\pi(p_1, f_1) \cup \exists_\pi(p_2, f_2) = p_1 \cup p_2 = 1_C;$$

but this says exactly that g is a cover. So it suffices to show that g is monic. Let $(k,l): F \to E$ be two morphisms with $gk = gl$. Since E is the union of A_1 and A_2, we get two different decompositions of F as a union of subobjects:

$$F = k^*(A_1) \cup k^*(A_2) = l^*(A_1) \cup l^*(A_2)\ .$$

By applying the distributive law (1.4.2) in $\mathrm{Sub}(F)$, we obtain a decomposition of F as the union of the four subobjects $k^*(A_i) \cap l^*(A_j)$ $(i,j = 1,2)$. Our aim is to show that the equalizer of hk and hl is the whole of F, by showing that it contains each of these four subobjects; this will suffice to prove the result, since the pair (g,h) is monic and we must then have $k = l$. We shall consider the subobject $k^*(A_1) \cap l^*(A_2)$ in detail; the other three are similar or easier. To simplify the notation, let us write G for $k^*(A_1) \cap l^*(A_2)$, and let $v_1 : G \to A_1$, $v_2 : G \to A_2$ be the unique morphisms making the squares

$$\begin{tikzcd}
A_1 \arrow[d, tail, "u_1"] & G \arrow[l, "v_1"'] \arrow[r, "v_2"] \arrow[d, tail, "w"] & A_2 \arrow[d, tail, "u_2"] \\
E & F \arrow[l, "k"] \arrow[r, "l"'] & E
\end{tikzcd}$$

commute. Now we have

$$p_1 v_1 = g u_1 v_1 = gkw = glw = g u_2 v_2 = p_2 v_2,$$

so the pair (v_1, v_2) factors through the pullback (n_1, n_2) (say by $x: G \to B$). But now $hu_1n_1 = f_1n_1 = f_2n_2 = hu_2n_2$, so we have

$$hkw = hu_1v_1 = hu_1n_1x = hu_2n_2x = hu_2v_2 = hlw,$$

and so the equalizer of hk and hl contains w, as required. □

Unions with the property that the square in the statement of 1.4.3 is a pushout have been called *effective unions* by Barr [78]. A particular case of 1.4.3 is of interest:

Corollary 1.4.4 *Let A_1 and A_2 be subobjects of an object A in a coherent category which are disjoint (i.e. such that $A_1 \cap A_2 \cong 0$). Then $A_1 \cup A_2$ is a coproduct of A_1 and A_2.*

Proof Pushouts under 0 are the same thing as coproducts. □

In a general category, a coproduct $A+B$ is called *disjoint* if the coprojections $A \to A+B$, $B \to A+B$ are monic and their intersection in $\mathrm{Sub}(A+B)$ is an initial object of the latter. We call a coherent category *positive* if it has disjoint finite coproducts; by 1.4.4, this is equivalent to saying that any two objects of the category may be embedded as disjoint subobjects of some third object. Incidentally, when working in a category where coproducts are known to be disjoint, we shall usually denote them by the disjoint union sign $\amalg$ rather than the addition sign $+$.

If $F: \mathcal{C} \to \mathcal{D}$ is a *coherent functor* between coherent categories (i.e. a regular functor which preserves finite unions) and $\mathcal{C}$ is positive, then F necessarily preserves finite coproducts as well. In general, however, even when a coherent category has finite coproducts they need not be disjoint, nor preserved by coherent functors. Any distributive lattice, considered as a preorder, is a coherent category (the distributivity is necessary, by 1.4.2), but it is not positive unless it is degenerate (i.e. satisfies $0 \cong 1$). If $\mathcal{C}$ is a coherent category which is not a preorder (e.g. $\mathcal{C} = \mathbf{Set}$), then the inclusion $\mathrm{Sub}(1) \to \mathcal{C}/1 \cong \mathcal{C}$ is a coherent functor which does not preserve coproducts.

The category $\mathbf{Set}$ is coherent and positive. If $\mathcal{C}$ is coherent, so are any functor category $[\mathcal{D}, \mathcal{C}]$ (unions of subobjects in $[\mathcal{D}, \mathcal{C}]$ are constructed pointwise; they are preserved by pullback because the latter is also computed pointwise) and any slice category $\mathcal{C}/B$ (we have an isomorphism $\mathrm{Sub}_{\mathcal{C}/B}(f) \cong \mathrm{Sub}_{\mathcal{C}}(\mathrm{dom}\ f)$, which commutes with pullback functors); and positivity is inherited by both these constructions. $\mathbf{LH}$ (like $\mathbf{Sp}$) has disjoint coproducts which are preserved by pullback along an arbitrary morphism; from this, and the remarks at the beginning of this section, it follows that $\mathbf{LH}/X$ is a positive coherent category for any X. However, categories which are monadic over $\mathbf{Set}$, like $\mathbf{Gp}$ and $\mathbf{Ab}$, tend not to be coherent; the two examples just cited have neither strict initial objects nor distributive subobject lattices. ($\mathbf{Ab}$ does have effective unions – as does any abelian category, for reasons entirely different from those adduced in the proof of 1.4.3 – but $\mathbf{Gp}$ does not: consider any two nontrivial subgroups of the nonabelian group of order 6.)

If $\mathcal{C}$ is a positive coherent category and $f: A \to \coprod_{i=1}^{m} B_i$ is a morphism of $\mathcal{C}$ whose codomain is an m-fold coproduct, then the subobjects $A_i = f^*(B_i)$ $(1 \leq i \leq m)$ of A form an *m-fold decomposition* of A; that is, they are pairwise disjoint, and their union is the whole of A. (We allow the possibility that some of the A_i may be 0.) Using this idea, we may describe a (non-elementary) construction for 'positivizing' a given coherent category by freely adjoining disjoint coproducts, as follows. Given $\mathcal{C}$, we define $\mathbf{Pos}(\mathcal{C})$ to be a category whose objects are nonempty finite sequences $(A_1, A_2, \ldots, A_n)$ of objects of $\mathcal{C}$ (we could allow the empty sequence as well, but since $\mathcal{C}$ already has a strict initial object we don't need to). Here $(A_1, \ldots, A_n)$ is to be thought of as the coproduct of the A_i in $\mathbf{Pos}(\mathcal{C})$, so we need only describe those morphisms of $\mathbf{Pos}(\mathcal{C})$ whose domains are sequences of length 1, say $f: (A) \to (B_1, \ldots, B_m)$; such an f is specified by an m-fold decomposition $(A_1, \ldots, A_m)$ of A and a family of morphisms $f_i: A_i \to B_i$

$(1 \leq i \leq m)$ in $\mathcal{C}$. Composition of morphisms in $\mathbf{Pos}(\mathcal{C})$ is defined in the obvious (but messy) way; we may regard $\mathcal{C}$, up to isomorphism, as a full subcategory of $\mathbf{Pos}(\mathcal{C})$, whose objects are the sequences of length 1. (Note also that if a finite family of objects $(A_1, \ldots, A_m)$ happens to have a disjoint coproduct in $\mathcal{C}$, then this sequence is (canonically) isomorphic in $\mathbf{Pos}(\mathcal{C})$ to the singleton sequence $(\coprod_{i=1}^{m} A_i)$.)

Proposition 1.4.5 $\mathbf{Pos}(\mathcal{C})$ *is a positive coherent category. Moreover, the embedding* $\mathcal{C} \to \mathbf{Pos}(\mathcal{C})$ *is a coherent functor, and it is universal (in the 'up-to-unique-isomorphism' sense) among coherent functors from* $\mathcal{C}$ *to positive coherent categories.*

Proof This is all straightforward but tedious verification, and we omit most of it. The product of two objects $(A_1, \ldots, A_n)$ and $(B_1, \ldots, B_m)$ is the sequence of length mn whose entries are the products $A_i \times B_j$; their coproduct is the sequence of length $m + n$ obtained by concatenating them. The image of a morphism $f\colon (A_1, \ldots, A_n) \to (B_1, \ldots, B_m)$ may be written in the form $(C_1, \ldots, C_m) \rightarrowtail (B_1, \ldots, B_m)$, where $C_j \in \mathrm{Sub}(B_j)$ is the union, for $1 \leq i \leq n$, of the images of the morphisms from appropriate subobjects of A_i to B_j; in particular, we note that any subobject of $(B_1, \ldots, B_m)$ is isomorphic to one of the form $(C_1, \ldots, C_m)$ where each $C_j \in \mathrm{Sub}(B_j)$. (We also obtain a criterion for a morphism to be a cover, which we may use to verify that covers are stable under pullback in $\mathbf{Pos}(\mathcal{C})$.) If $F\colon \mathcal{C} \to \mathcal{D}$ is a coherent functor and $\mathcal{D}$ is positive, we define an extension $\overline{F}\colon \mathbf{Pos}(\mathcal{C}) \to \mathcal{D}$ of F by setting

$$\overline{F}(A_1, \ldots, A_n) = \coprod_{i=1}^{n} F(A_i)$$

(and verify that this is a functor); but since any coherent functor $\mathbf{Pos}(\mathcal{C}) \to \mathcal{D}$ extending F must preserve finite coproducts, it is canonically naturally isomorphic to $\overline{F}$. □

Example 1.4.6 Let L be a distributive lattice, regarded as a coherent category. We may describe $\mathbf{Pos}(L)$ explicitly as follows: its objects are finite sequences $(l_1, \ldots, l_n)$ of elements of L, and a morphism $(k_1, \ldots, k_m) \to (l_1, \ldots, l_n)$ may be represented as an $n \times m$ matrix (f_{ij}) of elements of L, with the property that the jth column is an n-fold decomposition of k_j for each j, and the entries in the ith row are all below l_i, for each i. Composition of morphisms in this case is simply 'matrix multiplication', with addition and multiplication of matrix entries interpreted as join and meet of elements of L. We shall meet (close relatives of) this category again in 2.6.4(e) and 3.2.2(b); for the present, let us note that in the particular case when L is the two-element lattice $\mathbf{2} = \{0, 1\}$, we obtain a category equivalent to the category $\mathbf{Set}_f$ of finite sets and functions between them (since a matrix satisfying the conditions above defines a

function from $\{j \mid k_j = 1\}$ to $\{i \mid l_i = 1\}$, sending j to the unique i such that $f_{ij} = 1$).

In the particular case just described, $\mathbf{Pos}(L)$ is effective as a regular category as well as positive as a coherent category; but in general the positivization process can destroy effectiveness. Any lattice L is effective as a regular category (vacuously so, since reflexivity forces each equivalence relation in L to be a pair of identity maps), but $\mathbf{Pos}(L)$ is not effective unless L is a Boolean algebra. If l is any non-complemented element of L, then the two morphisms $(1, l, l, 1) \to (1, 1)$ in $\mathbf{Pos}(L)$ defined by the matrices

$$\begin{pmatrix} 1 & l & 0 & 0 \\ 0 & 0 & l & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 & l & 0 \\ 0 & l & 0 & 1 \end{pmatrix}$$

may be shown to define an equivalence relation which is not effective. Fortunately, however, the effectivization process which we shall describe in 3.3.10 does not destroy coherence or positivity.

From the fact that $\mathbf{Pos}(\mathbf{2}) \simeq \mathbf{Set}_f$, we may deduce a result which will be useful later.

Corollary 1.4.7 *For any positive coherent category $\mathcal{C}$, there is (up to isomorphism) a unique coherent functor $\Delta\colon \mathbf{Set}_f \to \mathcal{C}$.*

Proof There is clearly a unique coherent functor $\mathbf{2} \to \mathcal{C}$, since such functors must preserve 0 and 1; so this is immediate from 1.4.5. □

The functor Δ is easy to describe explicitly: it sends an n-element set to the n-fold copower of 1 in $\mathcal{C}$. It is worth noting that it is not merely coherent but cocartesian (that is, it preserves arbitrary coequalizers and not just those of equivalence relations); this is because the equivalence relation generated by a parallel pair in $\mathbf{Set}_f$ may be constructed from the latter using finite limits, finite coproducts and image factorizations, all of which are preserved by Δ.

A coherent category which is both positive and effective is called a *pretopos*. We note that a pretopos is 'almost cocartesian': it has finite coproducts and coequalizers of equivalence relations (and any coherent functor defined on it preserves these), but it may fail to have arbitrary coequalizers (and even if it does, a coherent functor defined on it may fail to preserve them – see D3.3.9 for a counterexample, but cf. also 1.4.19 below). A pretopos also has an important

class of pushouts:

Lemma 1.4.8 *In a pretopos, any pair of monomorphisms with common domain has a pushout. Moreover, if*

$$\begin{tikzcd} A \arrow[r, tail, "f"] \arrow[d, tail, "g"] & B \arrow[d, "h"] \\ C \arrow[r, "k"] & P \end{tikzcd}$$

is a pushout square with f and g monic, then h and k are monic, and the square is also a pullback.

Proof Given f and g, it suffices to find an object D containing B and C as subobjects in such a way that their intersection is A; for then we can define $P = B \cup C$ and appeal to Proposition 1.4.3. Let $D_0 = B \amalg C$, and let $u: B \to D_0$, $v: C \to D_0$ be the coprojections; define

$$R = B \amalg A_1 \amalg A_2 \amalg C$$

(where A_1 and A_2 are both copies of A, which we have tagged for future reference), and let $a, b: R \to D_0$ be the morphisms whose components are respectively (u, uf, vg, v) and (u, vg, uf, v). It is straightforward to verify that (a, b) is an equivalence relation; let $q: D_0 \to D$ be its coequalizer. Then (a, b) is the kernel-pair of q, and hence we may construct a diagram

$$\begin{tikzcd} A_1 \arrow[r, tail] \arrow[d, tail] & B \amalg A_1 \arrow[r, "{(1, f)}"] \arrow[d, tail] & B \arrow[d, tail, "u"] \\ A_1 \amalg C \arrow[r, tail] \arrow[d, "{(g, 1)}"] & R \arrow[r, "a"] \arrow[d, "b"] & D_0 \arrow[d, "q"] \\ C \arrow[r, tail, "v"] & D_0 \arrow[r, "q"] & D \end{tikzcd}$$

in which all squares are pullbacks (and the unnamed maps are coprojections). So we need only verify that qu and qv are monic; but we have a similar diagram

of pullbacks

$$\begin{tikzcd}
B \arrow[r, tail] \arrow[d, tail] & B \amalg A_1 \arrow[r, "{(1,f)}"] \arrow[d, tail] & B \arrow[d, tail, "u"] \\
B \amalg A_2 \arrow[r, tail] \arrow[d, "{(1,f)}"] & R \arrow[r, "a"] \arrow[d, "b"] & D_0 \arrow[d, "q"] \\
B \arrow[r, tail, "u"] & D_0 \arrow[r, "q"] & D
\end{tikzcd}$$

which establishes that qu is monic, and similarly for qv. □

Note that the object D constructed in the proof of 1.4.8, being covered by $B \amalg C$, is actually the union of its subobjects $qu \colon B \to D$ and $qv \colon C \to D$; thus it is itself the pushout P which we required.

Corollary 1.4.9 *In a pretopos, every monomorphism is regular (i.e. occurs as an equalizer) and every epimorphism is regular. In particular, a pretopos is a balanced category.*

Proof If $f \colon A \to B$ is monic, then by 1.4.8 we have a pullback square of the form

$$\begin{tikzcd}
A \arrow[r, tail, "f"] \arrow[d, tail, "f"] & B \arrow[d, tail, "g"] \\
B \arrow[r, tail, "h"] & C
\end{tikzcd}$$

from which it follows that f is an equalizer of g and h. The third assertion follows because an epic equalizer is necessarily an isomorphism; now if f is any epimorphism in $\mathcal{C}$, then im f is both epic and monic, so it is an isomorphism and f is a cover, hence regular epic by 1.3.4. □

A subobject $A' \rightarrowtail A$ in a coherent category is said to be *complemented* if there exists $A'' \rightarrowtail A$ such that $A' \cap A'' \cong 0$ and $A' \cup A'' \cong A$. By 1.4.2, the complement A'' is unique up to isomorphism if it exists. We say a coherent category is *Boolean* if every subobject in it has a complement; more precisely, if we are given an operation $\neg$ assigning to each monomorphism $A' \rightarrowtail A$ a complement $\neg A' \rightarrowtail A$ for it in $\mathrm{Sub}(A)$. Note that $\neg$ is automatically a functor $\mathrm{Sub}(A)^{\mathrm{op}} \to \mathrm{Sub}(A)$ (i.e. if $A' \leq A''$ then $\neg A'' \leq \neg A'$), and its square $\neg\neg$ is isomorphic to the identity. It is clear that pullback functors $f^* \colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$

preserve complements (i.e. they commute up to isomorphism with $\neg$), from which we deduce

Lemma 1.4.10 *Let $f\colon A \to B$ be a morphism in a Boolean coherent category. Then the functor $f^*\colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ has a right adjoint $\forall_f$, as well as its left adjoint $\exists_f$.*

Proof Define $\forall_f = \neg\exists_f\neg$: by the remarks above, this is right adjoint to $\neg f^*\neg$, which is isomorphic to f^*. □

A coherent category $\mathcal{C}$ satisfying the conclusion of Lemma 1.4.10 (for every morphism f of $\mathcal{C}$) is called a *Heyting category* (the Freyd–Scedrov name is 'logos'; we have chosen a name which emphasizes the fact that these categories are generalizations of Heyting algebras – see 1.5.11 below). A *Heyting functor* between Heyting categories is a coherent functor which commutes up to isomorphism with the right adjoints $\forall_f$. It follows easily from the proof of 1.4.10 that a coherent functor whose domain is Boolean must be a Heyting functor (even if its range is not Boolean). Further examples of Heyting functors are the pullback functors $f^*\colon \mathcal{C}/B \to \mathcal{C}/A$, where $f\colon A \to B$ is a morphism in a Heyting category $\mathcal{C}$; that is,

Lemma 1.4.11 *Let*

$$\begin{tikzcd} P \arrow[r, "h"] \arrow[d, "k"] & A \arrow[d, "f"] \\ B \arrow[r, "g"] & C \end{tikzcd}$$

be a pullback square in a Heyting category $\mathcal{C}$. Then the square

$$\begin{tikzcd} \mathrm{Sub}(A) \arrow[r, "h^*"] \arrow[d, "\forall_f"] & \mathrm{Sub}(P) \arrow[d, "\forall_k"] \\ \mathrm{Sub}(C) \arrow[r, "g^*"] & \mathrm{Sub}(B) \end{tikzcd}$$

commutes up to isomorphism.

Proof The corresponding diagram of left adjoints

$$\begin{tikzcd} \mathrm{Sub}(A) & \mathrm{Sub}(P) \arrow[l, "\exists_h"'] \\ \mathrm{Sub}(C) \arrow[u, "f^*"] & \mathrm{Sub}(B) \arrow[l, "\exists_g"'] \arrow[u, "k^*"] \end{tikzcd}$$

commutes, by the definition of the functors $\exists_{(-)}$ and of a regular category. So the result follows from uniqueness of adjoints. $\square$

The commutativity of the square in the statement of 1.4.11 (and of the corresponding square for existential quantification, appearing in the proof) is sometimes referred to as the *Beck–Chevalley condition.* We shall meet this condition again, in different contexts, in 2.2.4 and 4.1.16 (and yet again in B1.4.2).

The category **Set** is Boolean (and 'complement' means what it usually does). Booleanness is stable under slicing (and under the positivization process of 1.4.5), but not in general under passage to functor categories; in fact we have

Lemma 1.4.12 *For a locally small category $\mathcal{C}$, the functor category $[\mathcal{C}, \mathbf{Set}]$ is Boolean iff $\mathcal{C}$ is a groupoid.*

Proof Suppose $\mathcal{C}$ is a groupoid; let $F: \mathcal{C} \to \mathbf{Set}$ be a functor, and $F' \rightarrowtail F$ a subfunctor. We may suppose without loss of generality that the values $F'(A)$, $A \in \operatorname{ob} \mathcal{C}$, of F' are actual subsets of the corresponding $F(A)$. Define

$$F''(A) = \{x \in F(A) \mid x \notin F'(A)\};$$

then F'' is also a subfunctor of F, since if $f: A \to B$ in $\mathcal{C}$ and $x \in F''(A)$ then we cannot have $F(f)(x) \in F'(B)$, for else $x = F(f^{-1})F(f)(x)$ would be in $F'(A)$. And F'' is clearly a complement for F', since unions and intersections of subobjects in $[\mathcal{C}, \mathbf{Set}]$ are computed pointwise.

Conversely, suppose $[\mathcal{C}, \mathbf{Set}]$ is Boolean. The functor $F = \mathcal{C}(A, -)$ has no complemented subobjects other than itself and 0, for if $F = F' \cup F''$ then one of $F'(A)$ and $F''(A)$ must contain 1_A, which forces the corresponding one of F' and F'' to be the whole of F. Thus if $f: A \to B$ is any morphism of $\mathcal{C}$, the subfunctor F_f of F defined by

$$F_f(C) = \{g: A \to C \mid g \text{ factors through } f\}$$

must be the whole of F (since it is clearly nonzero); so $1_A \in F_f(A)$, i.e. f is a split monomorphism. But a category in which every morphism is split monic is a groupoid. $\square$

Subobjects of X in **LH** correspond (up to isomorphism) to open subsets of X; so if $\mathbf{LH}/X$ is Boolean then all its open subsets must be closed. If X satisfies the T_0 separation axiom, this forces it to be discrete; but for a discrete space X, we have an isomorphism $\mathbf{LH}/X \cong \mathbf{Set}/|X|$ (where $|X|$ is, as usual, the underlying set of X), and so $\mathbf{LH}/X$ is Boolean in this case.

Despite the foregoing results, $[\mathcal{C}, \mathbf{Set}]$ is a Heyting category for any small category $\mathcal{C}$, and $\mathbf{LH}/X$ is a Heyting category for any space X. Both assertions will follow from 1.4.18 below (and the former will also follow from two results to be proved in the next section, namely 1.5.5 and 1.5.13).

In a coherent category, we define the *negation* (or pseudo-complement) of a subobject $A' \rightarrowtail A$ to be the largest subobject of A disjoint from A', if this exists. We shall denote it by $\neg A' \rightarrowtail A$; there is no harm in re-using our notation for complements, since it is clear that a complement (if it exists) is a pseudo-complement, and conversely a pseudo-complemented subobject $A' \rightarrowtail A$ is complemented iff $A' \cup \neg A' \cong A$. In a Heyting category, negations always exist: indeed, we have a more general construction.

Lemma 1.4.13 *Let $A_1 \rightarrowtail A$ and $A_2 \rightarrowtail A$ be subobjects in a Heyting category. Then there exists a largest subobject $A_3 \rightarrowtail A$ such that $A_3 \cap A_1 \leq A_2$. Moreover, if we denote this subobject by $(A_1 \Rightarrow A_2) \rightarrowtail A$, then the binary operation on subobjects thus defined is stable under pullback.*

Proof We define $(A_1 \Rightarrow A_2) \rightarrowtail A$ to be $\forall_m (A_1 \cap A_2 \rightarrowtail A_1)$, where m is the name of the monomorphism $A_1 \rightarrowtail A$. It is an immediate consequence of the adjunction $(m^* \dashv \forall_m)$ that this subobject has the required property; the stability under pullback follows from 1.4.11. $\square$

The stability of $\Rightarrow$ (when it exists) under pullback is no surprise, since this assertion is equivalent to the Frobenius reciprocity condition established for regular categories in 1.3.3 – see 1.5.8 below. In the terminology to be introduced in the next section, 1.4.13 says that subobject lattices in Heyting categories are *Heyting algebras*. Of course, we obtain the negation $\neg A'$ as $(A' \Rightarrow 0)$. For future reference, we note

Lemma 1.4.14 *A Heyting category $\mathcal{C}$ is Boolean iff $\neg\neg \colon \mathrm{Sub}(A) \to \mathrm{Sub}(A)$ is (isomorphic to) the identity, for every object A of $\mathcal{C}$.*

Proof The necessity of the condition has already been mentioned. For the sufficiency, observe that we always have

$$\neg(A' \cup \neg A') \cong 0,$$

because a subobject A'' disjoint from $A' \cup \neg A'$ is in particular disjoint from A', and hence contained in $\neg A'$. So

$$\neg\neg(A' \cup \neg A') \cong A,$$

from which the result follows. $\square$

An object A of a coherent category is said to be *decidable* if the diagonal map $\Delta_A \colon A \rightarrowtail A \times A$ is a complemented subobject. Of course, in a Boolean coherent category every object is decidable; but the converse is false. In a distributive lattice considered as a coherent category, every object is decidable since diagonal maps are isomorphisms; but we can also give positive counterexamples using 1.4.15 and 1.4.16 below. (However, in a topos (and, more generally, in a coherent category with a subobject classifier, as defined in Section A1.6 below), the

decidability of every object is equivalent to Booleanness. This is because, if A is decidable, then every morphism $1 \to A$ is a complemented subobject (since it is split monic, and so can be expressed as an equalizer of 1_A and $A \to 1 \to A$, or equivalently as a pullback of Δ_A), and so decidability of the subobject classifier Ω forces the generic subobject $\top \colon 1 \rightarrowtail \Omega$ to be complemented.)

We note that the objects 0 and 1 are always decidable (their diagonal maps are isomorphisms); any subobject of a decidable object is decidable (since, if $f \colon A \rightarrowtail B$ is monic, then Δ_A is the pullback of Δ_B along $f \times f$); a disjoint coproduct $A \amalg B$ of decidable objects is decidable (since we have

$$(A \amalg B) \times (A \amalg B) \cong (A \times A) \amalg (A \times B) \amalg (B \times A) \amalg (B \times B)$$

by stability of unions under pullback, and the complement of the diagonal $\Delta_{A \amalg B}$ is then the union of the second and third summands above with the complements of Δ_A and Δ_B); and a product $A \times B$ of decidable objects is decidable (since its diagonal is the intersection of the pullbacks of Δ_A and Δ_B along the appropriate product projections, and a finite intersection of complemented subobjects is complemented). Thus we have proved

Proposition 1.4.15 *If $\mathcal{C}$ is a coherent category, then the full subcategory $\mathcal{C}_d$ of decidable objects of $\mathcal{C}$ is also coherent, and positive if $\mathcal{C}$ is; and the inclusion $\mathcal{C}_d \to \mathcal{C}$ is a coherent functor.* □

Examples 1.4.16 (*a*) In a functor category $[\mathcal{C}, \mathbf{Set}]$, an object F is decidable iff the assignment

$$A \mapsto \{(x, y) \in FA \times FA \mid x \neq y\}$$

defines a subfunctor of $F \times F$, iff $F(f)$ is injective for every morphism f of $\mathcal{C}$. Using this, it is easy to give examples of categories $\mathcal{C}$ for which $[\mathcal{C}, \mathbf{Set}]_d$ is not Boolean.

(*b*) In the category $\mathbf{LH}/X$, an object $p \colon E \to X$ is decidable iff the diagonal embedding $E \rightarrowtail E \times_X E$ is closed as well as open, since subobjects in $\mathbf{LH}$ correspond to open subspaces. If X is a Hausdorff space (i.e. if the diagonal $X \rightarrowtail X \times X$ is closed), this condition is equivalent to saying that E is also Hausdorff, since the canonical inclusion $E \times_X E \rightarrowtail E \times E$ is a pullback of $X \rightarrowtail X \times X$ and hence closed.

We remark that, although decidability is preserved by coherent functors, it is not reflected by conservative coherent functors – as may be seen by considering the forgetful functor $[\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}^{\mathrm{ob}\,\mathcal{C}}$ for suitable $\mathcal{C}$.

Although, in this chapter, our main concern is with the elementary (and hence finitary) structure of categories, it is worth digressing briefly, before we end this section, to consider infinite unions and intersections.

In the context of a particular model of set theory, we shall say that a category $\mathcal{C}$ is *well-powered* if each of the preorders $\mathrm{Sub}(A)$, $A \in \mathrm{ob}\,\mathcal{C}$, is equivalent to a small category. In passing, we remark that category-theorists often

refer to the dual condition (i.e. the well-poweredness of $\mathcal{C}^{\mathrm{op}}$) by saying '$\mathcal{C}$ is cowell-powered'; the rules of syntax clearly demand that the term should be *well-copowered*, and that is what we shall use in this book.

Remark 1.4.17 The property of being well-powered is closely related to local smallness. First, a well-powered category with finite products is locally small, since if two morphisms $A \rightrightarrows B$ are not equal then their graphs are non-isomorphic subobjects of $A \times B$ (cf. 2.2.3 below). Conversely, a locally small category with finite intersections of subobjects and a generating set of objects is well-powered, since a subobject of a given object A is determined (up to isomorphism) by the set of morphisms from members of the generating set to A which factor through it. And any locally small category with a subobject classifier as defined in Section A1.6 below (in particular, any locally small topos) is well-powered, even if it does not have a generating set.

We note that the particular category $\mathbf{Reg}(\mathcal{C})$ which we considered in 1.3.10(d) is locally small but not well-powered; hence it does not have a generating set. On the other hand, a 'large group' (that is, a groupoid with one object but a proper class of morphisms) is well-powered (since any two subobjects are isomorphic) but not locally small.

When talking about well-powered categories, we generally interpret 'subobject' as meaning 'isomorphism class of monomorphisms', so that $\mathrm{Sub}(A)$ becomes a (small) partial order rather than a preorder. Of course, such an ordered set has arbitrary (set-indexed) unions iff it has arbitrary intersections; the latter are automatically preserved by pullback functors (if $\mathcal{C}$ has pullbacks) since they have left adjoints, but the former may not be. In fact we have

Lemma 1.4.18 *Let $\mathcal{C}$ be a well-powered regular category whose subobject lattices have arbitrary (unions and) intersections. Then $\mathcal{C}$ is a Heyting category iff arbitrary unions of subobjects are stable under pullback in $\mathcal{C}$.*

Proof $f^* \colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ preserves arbitrary unions iff it has a right adjoint, by the poset version of the adjoint functor theorem. □

A category satisfying the conditions of Lemma 1.4.18 will be called a *geometric category*. As hinted above, both $[\mathcal{C}, \mathbf{Set}]$ (for any small $\mathcal{C}$) and $\mathbf{LH}/X$ (for any space X) are geometric categories; in the first case this follows from the fact that both unions and pullbacks are computed pointwise (and $\mathbf{Set}$ is a geometric category), and in the second from the fact that subobjects correspond to open subsets.

We shall say that a geometric category $\mathcal{C}$ is ∞-*positive* if every set-indexed family of objects of $\mathcal{C}$ has a disjoint coproduct (equivalently, can be embedded as a pairwise-disjoint family of subobjects of a single object). The category $\mathbf{Set}_f$ is an example of a geometric category (since its subobject lattices are finite) which is positive but not ∞-positive. Of course, there is an infinitary analogue of the positivization process of 1.4.5, in which we take the objects of $\mathbf{Pos}_\infty(\mathcal{C})$ to be

arbitrary set-indexed families of objects of $\mathcal{C}$, rather than finite families. And we say $\mathcal{C}$ is an ∞-*pretopos* if it is an ∞-positive geometric category which is effective as a regular category. For future reference, we note

Lemma 1.4.19 *An ∞-pretopos is cocomplete.*

Proof Since an ∞-pretopos $\mathcal{C}$ has set-indexed coproducts by definition, it suffices to show that it has coequalizers. But by 1.3.4 we know that equivalence relations in $\mathcal{C}$ have coequalizers; so it suffices to show that, for any parallel pair $f, g: B \rightrightarrows A$ in $\mathcal{C}$, there is an equivalence relation $a, b: R \rightrightarrows A$ such that a morphism $h: A \to C$ satisfies $hf = hg$ iff $ha = hb$.

First we form the image factorization

$$B \xrightarrow{\quad q \quad} I \rightarrowtail^{(c,d)} A \times A$$

of (f, g); since q is epic, we have $hf = hg$ iff $hc = hd$. Next we replace I by its reflexive and symmetric closure $\overline{I} = I \cup \Delta \cup I^{\circ}$ in $\mathrm{Sub}(A \times A)$; using the fact that unions are pushouts (1.4.3), it is clear that any morphism which coequalizes c and d will also coequalize the morphisms $\overline{c}, \overline{d}: \overline{I} \rightrightarrows A$. Then we construct the transitive closure R of $\overline{I}$, as the union of its finite powers (each of which is constructed as the image of a map to $A \times A$ from an appropriate pullback); yet again, a morphism which coequalizes $\overline{c}$ and $\overline{d}$ will coequalize the morphisms from each of these finite powers to A, and hence (by the infinitary analogue of 1.4.3, which is proved in the same way as the original) it will coequalize the projections $a, b: R \rightrightarrows A$. But R is an equivalence relation; so this is sufficient. □

We shall also occasionally wish to consider regular categories which have pullback-stable (finite or) countable unions of subobjects, but not necessarily unions of larger cardinality. We shall call these σ-*coherent categories*, or σ-*pretoposes* if they are effective and σ-positive (i.e. have countable disjoint coproducts). We note that the proof of 1.4.19 uses only countable unions, and so shows that a σ-pretopos has all countable colimits.

Suggestions for further reading: Barr [78], Börger [169], Freyd & Scedrov [381], Ghilardi & Zawadowski [403], Makkai & Reyes [790].

A1.5 Cartesian closed categories

The next level of 'set-like' categorical structure which we study is something of a new departure: it is concerned not with the preorders $\mathrm{Sub}(A)$, $A \in \mathrm{ob}\,\mathcal{C}$, but with the whole category $\mathcal{C}$. To put the same thing another way: the classes of categories studied in the last three sections correspond (in a sense made precise in Chapter D1) to certain fragments of first-order logic, but cartesian closed categories correspond to the typed lambda-calculus (see Section D4.2), which belongs to the realm of higher-order logic.

Let $\mathcal{C}$ be a category with finite products. For a fixed object A of $\mathcal{C}$, we can regard the assignment $B \mapsto B \times A$ as the object-map of a functor $(-) \times A : \mathcal{C} \to \mathcal{C}$. We say that A is *exponentiable* if this functor has a right adjoint (generally denoted $(-)^A$, though we may use other notations such as $[A, -]$ or $(A \Rightarrow -)$ from time to time). In elementary terms, this adjunction is best described in terms of its counit: A is exponentiable if we are given an operation assigning to each object B of $\mathcal{C}$ an object B^A and a morphism $\mathrm{ev}: B^A \times A \to B$ (called 'evaluation') such that for any $h: C \times A \to B$ there is a unique $\bar{h}: C \to B^A$ for which

$$\begin{tikzcd} C \times A \arrow[r, "\bar{h} \times 1"] \arrow[dr, "h"'] & B^A \times A \arrow[d, "\mathrm{ev}"] \\ & B \end{tikzcd}$$

commutes. We call $\bar{h}$ the *exponential transpose* of h.

Let $\mathcal{C}_{\mathrm{exp}}$ denote the full subcategory of exponentiable objects of $\mathcal{C}$.

Lemma 1.5.1

(i) *$\mathcal{C}_{\mathrm{exp}}$ has finite products, and the inclusion $\mathcal{C}_{\mathrm{exp}} \to \mathcal{C}$ preserves them.*

(ii) *The assignment $(A, B) \mapsto B^A$ defines a bifunctor $(\mathcal{C}_{\mathrm{exp}})^{\mathrm{op}} \times \mathcal{C} \to \mathcal{C}$.*

(iii) *If $\mathcal{C}$ is Cauchy-complete, then so is $\mathcal{C}_{\mathrm{exp}}$.*

Proof (i) The terminal object 1 is exponentiable, because $(-) \times 1$ is isomorphic to the identity functor. If A and B are exponentiable, then the object $(C^B)^A$, together with the composite

$$(C^B)^A \times A \times B \xrightarrow{\mathrm{ev} \times 1} C^B \times B \xrightarrow{\mathrm{ev}} C,$$

has the universal property of an exponential $C^{(A \times B)}$; so $A \times B$ is exponentiable.

(ii) Given $f: A' \to A$ and $g: B \to B'$, we define $g^f : B^A \to B'^{A'}$ to be the exponential transpose of

$$B^A \times A' \xrightarrow{g^A \times f} B'^A \times A \xrightarrow{\mathrm{ev}} B';$$

it is straightforward to verify that this works.

(iii) Suppose A is exponentiable, and let $e: A \to A$ be an idempotent endomorphism. Then we know e splits in A as $A \twoheadrightarrow B \rightarrowtail A$; we have to show that B is also exponentiable. But, for any C, we have an idempotent $C^e : C^A \to C^A$ by part (ii); and it is straightforward to verify that the image of this idempotent has the universal property of an exponential C^B. □

In **Gp** and **Ab**, the only exponentiable object is the terminal object. In **Sp**, any locally compact space is exponentiable (and the converse is true, modulo the need to take some care in defining 'locally compact'), the exponential Y^X being the set of continuous maps $X \to Y$ equipped with the compact-open topology. In **Set**, and in the category **Cat** of small categories, every object is exponentiable: in **Set** we take the exponential B^A to be the set of functions $A \to B$, and in **Cat** we take $\mathcal{D}^{\mathcal{C}}$ to be the functor category $[\mathcal{C}, \mathcal{D}]$. It is the latter situation we wish to study in this section.

We say a category $\mathcal{C}$ is *cartesian closed* if it has finite products and all its objects are exponentiable. This terminology is very well established, and we do not feel able to depart from it; but the reader must be warned that it produces a slight awkwardness, in that a cartesian closed category need not be cartesian in the sense of having all finite limits. (The primary justification for this is provided by the connection with the typed lambda-calculus, to which we have already alluded; the cartesian closed categories arising from models of the latter do not in general have limits other than finite products. However, another justification may be found in the fact that, even when $\mathcal{C}$ is cartesian, the subcategory $\mathcal{C}_{\mathrm{exp}}$ is not in general closed under equalizers, in contrast to 1.5.1(i) and (iii) – **Sp** provides a counterexample. See also 1.5.4 below.) Nevertheless, most of the cartesian closed categories we shall meet will have equalizers as well as finite products; we shall indicate this (with a note of mild reproof in our voice) by calling them *properly cartesian closed.*

If $F: \mathcal{C} \to \mathcal{D}$ is a finite-product-preserving functor between cartesian closed categories, then we have morphisms $\theta_{A,B}: F(B^A) \to F(B)^{F(A)}$ for objects A, B of $\mathcal{C}$, which form a natural transformation between bifunctors $\mathcal{C}^{\mathrm{op}} \times \mathcal{C} \to \mathcal{D}$ (cf. 1.5.1(ii)): explicitly, $\theta_{A,B}$ is the transpose of

$$F(B^A) \times F(A) \cong F(B^A \times A) \xrightarrow{F(\mathrm{ev})} F(B)\,.$$

We say F is a *cartesian closed functor* if $\theta_{A,B}$ is an isomorphism for all $A, B \in \mathrm{ob}\,\mathcal{C}$ (with the same convention as before about 'properly cartesian closed').

Recall that if B is an object of a cartesian category $\mathcal{C}$, we use B also as the name of the unique morphism $B \to 1$, so that $B^*: \mathcal{C} \to \mathcal{C}/B$ denotes the right adjoint of the forgetful functor $\Sigma_B: \mathcal{C}/B \to \mathcal{C}$. (Explicitly, $B^*(A)$ is the product projection $A \times B \to B$.)

Lemma 1.5.2 *Let $\mathcal{C}$ be a cartesian category.*

(i) *An object B of $\mathcal{C}$ is exponentiable iff B^* has a right adjoint $\Pi_B: \mathcal{C}/B \to \mathcal{C}$.*

(ii) *If A is exponentiable in $\mathcal{C}$, then $B^*(A)$ is exponentiable in $\mathcal{C}/B$ for any B. Moreover, B^* preserves any exponentials which exist in $\mathcal{C}$.*

Proof (i) $(-) \times B$ is the composite

$$\mathcal{C} \xrightarrow{B^*} \mathcal{C}/B \xrightarrow{\Sigma_B} \mathcal{C},$$

so if Π_B exists we may define $(-)^B$ as the composite $\Pi_B \circ B^*$. Conversely, suppose B is exponentiable. Given $f: A \to B$, form the pullback

$$\begin{tikzcd} \Pi_B(f) \arrow[r] \arrow[d] & A^B \arrow[d, "f^B"] \\ 1 \arrow[r, "\overline{q}"] & B^B \end{tikzcd}$$

where q is the (isomorphic) projection $1 \times B \to B$. Now, for any object C of $\mathcal{C}$, morphisms $C \to \Pi_B(f)$ correspond to morphisms $\overline{h}: C \to A^B$ such that $f^B\overline{h} = \overline{q}C$, and hence to morphisms $h: C \times B \to A$ such that fh is the product projection $C \times B \to B$, i.e. to morphisms $B^*(C) \to f$ in $\mathcal{C}/B$. It is straightforward to verify that this correspondence is natural in C and f.

(ii) Given an object $f: C \to B$ of $\mathcal{C}/B$, define $f^{B^*(A)}$ to be the left vertical morphism in the pullback square

$$\begin{tikzcd} P \arrow[r] \arrow[d] & C^A \arrow[d, "f^A"] \\ B \arrow[r, "\overline{p}"] & B^A \end{tikzcd}$$

where p is the projection $B \times A \to B$. If $g: D \to B$ is another object of $\mathcal{C}/B$, then the product $g \times B^*(A)$ in $\mathcal{C}/B$ is the diagonal of the pullback square

$$\begin{tikzcd} D \times A \arrow[r, "g \times 1"] \arrow[d] & B \times A \arrow[d, "p"] \\ D \arrow[r, "g"] & B \end{tikzcd} ,$$

from which it follows easily that morphisms $h: g \times B^*(A) \to f$ in $\mathcal{C}/B$ correspond to morphisms $\overline{h}: D \to C^A$ in $\mathcal{C}$ such that $f^A\overline{h} = \overline{p}g$, i.e. to morphisms $g \to f^{B^*(A)}$ in $\mathcal{C}/B$. Once again, this correspondence is easily seen to be natural. Finally, if f is itself of the form $B^*(E)$, we see that the pullback P above may be rewritten in the form

$$\begin{tikzcd} E^A \times B \arrow[r, "1 \times \overline{p}"] \arrow[d] & E^A \times B^A \arrow[d] \\ B \arrow[r, "\overline{p}"] & B^A \end{tikzcd}$$

(where the vertical arrows are product projections), because $(-)^A$, being a right adjoint, preserves products. □

Corollary 1.5.3 *A cartesian category $\mathcal{C}$ is locally cartesian closed (i.e. each slice category $\mathcal{C}/B$ is cartesian closed) iff $f^* : \mathcal{C}/B \to \mathcal{C}/A$ has a right adjoint Π_f for each morphism $f: A \to B$ of $\mathcal{C}$. Moreover, if $\mathcal{C}$ is locally cartesian closed, then the functors f^* are cartesian closed.*

Proof Recall the isomorphism $(\mathcal{C}/B)/f \cong \mathcal{C}/A$. □

Unlike the other conditions we have considered so far (and despite the result of 1.5.2(ii)), cartesian closedness is not stable under slicing, so 'locally cartesian closed' is a stronger condition than 'cartesian closed'. (By convention, a locally cartesian closed category is assumed to have a terminal object, so that it is in particular cartesian closed – and indeed properly cartesian closed, since all its slice categories are required to have finite products, i.e. the category itself is required to have pullbacks.)

For example, **Cat** is not locally cartesian closed; this may be seen, given the result of 1.5.3, by observing that pullback functors in **Cat** do not in general preserve coequalizers. (Indeed, **Cat** is not even regular: there is a regular epimorphism $\mathbf{2} + \mathbf{2} \to \mathbf{3}$ in **Cat**, where $\mathbf{n}$ denotes an n-element totally ordered set, whose pullback along a suitable morphism $\mathbf{2} \to \mathbf{3}$ is not epic.) J. Giraud [406], and independently F. Conduché [247], determined the exponentiable objects in $\mathbf{Cat}/\mathcal{C}$ for an arbitrary $\mathcal{C}$; they are functors $F: \mathcal{D} \to \mathcal{C}$ such that, given a factorization $F(f) = gh$ of a morphism in the image of F, it can be lifted to a factorization $f = g'h'$ with $F(g') = g$ and $F(h') = h$, 'uniquely up to morphisms which are sent by F to identities'.

Remark 1.5.4 Despite the foregoing, it follows easily from 1.5.2(ii) that, if $\mathcal{C}$ is cartesian closed and B is any object of $\mathcal{C}$, then the full subcategory $\mathcal{C}_B$ of $\mathcal{C}/B$ whose objects are those of the form B^*A, $A \in \text{ob}\,\mathcal{C}$, is cartesian closed, though not in general properly so (it is closed under products in $\mathcal{C}/B$ because B^* preserves them, but not in general under equalizers). Equivalently, $\mathcal{C}_B$ may be regarded as the Kleisli category of the comonad on $\mathcal{C}$ whose functor part is $(-)\times B$, with counit and comultiplication given by the projection $\pi_1 : A \times B \to A$ and $1 \times \Delta: A \times B \to A \times B \times B$ respectively; of course, the Eilenberg–Moore category of this comonad (i.e., the category of coalgebras) is (isomorphic to) $\mathcal{C}/B$ itself. In 4.2.1 below we shall investigate when the category of coalgebras for a comonad on $\mathcal{C}$ inherits (local) cartesian closedness from $\mathcal{C}$; and we shall meet the Kleisli category $\mathcal{C}_B$ again in D4.2.10.

The category **Set** is locally cartesian closed; this may be seen using the equivalence $\mathbf{Set}/B \simeq \mathbf{Set}^B$ of 1.1.6, plus the (obvious) fact that a product of cartesian closed categories is cartesian closed. More generally,

Proposition 1.5.5 *For any small category $\mathcal{C}$, the functor category $[\mathcal{C}, \mathbf{Set}]$ is locally cartesian closed.*

Proof Since, by 1.1.7, any slice of $[\mathcal{C}, \mathbf{Set}]$ is equivalent to another category of the same type, it suffices to prove that $[\mathcal{C}, \mathbf{Set}]$ is cartesian closed. In fact the definition of exponentials in this category is forced upon us by the Yoneda lemma: if the exponential G^F exists, then elements of $G^F(A)$ must correspond bijectively to morphisms $\mathcal{C}(A, -) \to G^F$, and hence to morphisms $\mathcal{C}(A, -) \times F \to G$ in $[\mathcal{C}, \mathbf{Set}]$. But $[\mathcal{C}, \mathbf{Set}]$ is locally small, so we may take the set of all such morphisms as the definition of $G^F(A)$ (and make it into a functor $\mathcal{C} \to \mathbf{Set}$: if $f: A \to B$ in $\mathcal{C}$, $G^F(f)$ is the operation 'compose with $\mathcal{C}(f, -) \times 1: \mathcal{C}(B, -) \times F \to \mathcal{C}(A, -) \times F$'). We define the evaluation map $\mathrm{ev}: G^F \times F \to G$ by $\mathrm{ev}_A(\phi, x) = \phi_A(1_A, x)$, where $\phi: \mathcal{C}(A, -) \times F \to G$ and $x \in F(A)$. And, given a morphism $\theta: H \times F \to G$, we define its exponential transpose $\overline{\theta}: H \to G^F$ by

$$(\overline{\theta}_A(z))_B(f, x) = \theta_B(H(f)(z), x)$$

for $z \in H(A)$, $f: A \to B$ and $x \in F(B)$. The rest of the proof is straightforward verification. □

An alternative, less explicit, proof of 1.5.5 may be given by observing that $[\mathcal{C}, \mathbf{Set}]$ (and its slice categories) have all small colimits, and that pullback functors preserve them (because they do so in $\mathbf{Set}$, and both pullbacks and colimits in $[\mathcal{C}, \mathbf{Set}]$ are constructed pointwise), and then applying the Adjoint Functor Theorem to obtain the right adjoints needed for 1.5.3. However, the explicit description of exponentials in $[\mathcal{C}, \mathbf{Set}]$ has many uses: for example,

Scholium 1.5.6

(i) *Let $\mathcal{C}$ be a small category with finite products. Then the Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ preserves any exponentials which exist in $\mathcal{C}$. In particular, if $\mathcal{C}$ is cartesian closed, then the Yoneda embedding is a cartesian closed functor.*

(ii) *For any small category $\mathcal{C}$, the functor $\Delta: \mathbf{Set} \to [\mathcal{C}, \mathbf{Set}]$ which sends a set A to the constant functor with value A is cartesian closed.*

Proof (i) Let us write Y for the Yoneda embedding; recall that it sends an object A to the functor $\mathcal{C}(-, A)$. Now if A, B, C are three objects of $\mathcal{C}$, then elements of $Y(B)^{Y(A)}(C)$ are by definition morphisms $Y(C) \times Y(A) \to Y(B)$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; but these correspond bijectively to morphisms $C \times A \to B$ in $\mathcal{C}$ (since Y is full and faithful, and preserves products), and hence (if the exponential B^A exists in $\mathcal{C}$) to morphisms $C \to B^A$ – that is, to elements of $Y(B^A)(C)$. It is easily verified that these bijections define a natural isomorphism $Y(B^A) \cong Y(B)^{Y(A)}$, and that this isomorphism is the canonical morphism $\theta_{A,B}$ defined earlier.

(ii) Let S and T be two sets, and A an object of $\mathcal{C}$. Then 1.5.5 tells us that $\Delta T^{\Delta S}(A)$ is the set of natural transformations $\alpha: \mathcal{C}(A, -) \times \Delta S \to \Delta T$. Given such a transformation, we get a mapping $\alpha_0: S \to T$ defined by $\alpha_0(s) = \alpha_A(1_A, s)$. But then, by naturality, for any $f: A \to B$ we have

$\alpha_B(f, s) = \Delta T(f)(\alpha_A(1_A, s)) = \alpha_0(s)$; so the entire natural transformation is determined by α_0. Thus we have a bijection from the exponential T^S in **Set** to $\Delta T^{\Delta S}(A)$, for each A; it is straightforward to check that these bijections themselves form a natural isomorphism $\Delta(T^S) \to \Delta T^{\Delta S}$, and that this isomorphism is the canonical map $\theta_{S,T}$. □

In particular, 1.5.6(i) shows that any small cartesian closed category admits a full and faithful cartesian closed functor to a properly cartesian closed category.

The category $\mathbf{LH}/X$ is also cartesian closed for any space X (and hence locally cartesian closed); the proof of this fact is deferred until 4.2.4(e).

Exponentiability is closely related to the notion of partial product, which was originally introduced in the category **Sp** by B.A. Pasynkov [938] and extended to more general categories by R. Dyckhoff and W. Tholen [321]. In a cartesian category $\mathcal{C}$, a *partial product* of an object A and a morphism $f: E \to B$ is defined to be an object $P(A, f)$ equipped with morphisms $\rho: P(A, f) \to B$ and $\lambda: P(A, f) \times_B E \to A$, which is universal among such; i.e., given any object C equipped with morphisms $C \to B$ and $C \times_B E \to A$, there exists a unique morphism $C \to P(A, f)$ commuting (in the obvious sense) with the structure morphisms. We observe that the notion of partial product includes as special cases that of an ordinary categorical product $A \times B$ (obtained by taking $f = 1_B$) and that of an exponential A^E (obtained by taking $B = 1$).

The fundamental existence theorem for partial products, due to Dyckhoff and Tholen, is as follows.

Proposition 1.5.7 *Let $f: E \to B$ be a morphism in a cartesian category $\mathcal{C}$. Then the partial product $P(A, f)$ exists for all objects A of $\mathcal{C}$ iff f is exponentiable as an object of $\mathcal{C}/B$.*

Proof First suppose f is exponentiable. Then we define $\rho: P(A, f) \to B$ to be the exponential $(B^*A)^f$ in $\mathcal{C}/B$, and $\lambda: P(a, f) \times_B E = \Sigma_B(\rho \times f) \to A$ to be the transpose of the evaluation map $\rho \times f \to B^*A$. It is easy to verify that this has the right universal property: for, given an object C equipped with maps $r: C \to B$ and $l: C \times_B E \to A$, we can regard the latter as a morphism $\Sigma_B(r \times f) \to A$, and transpose it across the adjunctions to obtain the required morphism $r \to \rho$.

Conversely, if $P(A, f)$ exists for all A, then it is easily seen that $P(-, f)$ is (covariantly) functorial: given $h: A \to A'$, we get a morphism $P(A, f) \to P(A', f)$ induced by $\rho: P(A, f) \to B$ and $h\lambda: P(A, f) \times_B E \to A \to A'$. So, given any $g: A \to B$, we may define g^f to be the left vertical morphism in the pullback square

$$\begin{tikzcd} C \arrow[r] \arrow[d] & P(A, f) \arrow[d, "P(g, f)"] \\ B \arrow[r, "\overline{f}"] & P(B, f) \end{tikzcd}$$

where $\bar{f}$ corresponds to the pair $(1_B \colon B \to B, f \colon B \times_B E \cong E \to B)$. Then it is easily verified that, for any $h \colon D \to B$, morphisms $D \to C$ over B correspond to morphisms $k \colon D \to P(A, f)$ satisfying $P(g, f)k = \bar{f}h$, and these in turn correspond to morphisms $D \times_B E \to A$ over B. □

We shall see an application of partial products in 2.5.17 below. And in Section B4.4 we shall investigate a 2-categorical generalization of the notion of partial product, and show that significant examples of such partial products exist in the 2-category of toposes and geometric morphisms.

The following result concerning cartesian closed functors will be needed at numerous points later on.

Lemma 1.5.8 *Let $F \colon \mathcal{C} \to \mathcal{D}$ be a functor between cartesian closed categories, and suppose F has a left adjoint L. Then F is cartesian closed iff the canonical morphism*

$$(L\pi_1, \epsilon_A L\pi_2) \colon L(B \times FA) \longrightarrow LB \times A$$

(where ϵ is the counit of $(L \dashv F)$) is an isomorphism for all objects A, B of $\mathcal{C}$ and $\mathcal{D}$ respectively.

The isomorphism $L(B \times FA) \cong LB \times A$ is commonly called the *Frobenius reciprocity law* for the adjunction $(L \dashv F)$. Note that Lemmas 1.3.3 and 1.4.13 provide an example of an adjunction (between cartesian closed posets) satisfying the above conditions.

Proof First suppose F is cartesian closed. We construct an inverse for $(L\pi_1, \epsilon_A L\pi_2)$ by forming the composite

$$\begin{tikzcd}
LB \times A \arrow[r, "L\lambda \times 1"] & L((B \times FA)^{FA}) \times A \arrow[r, "L(\eta^{FA}) \times 1"] & L(FL(B \times FA)^{FA}) \times A \arrow[d, "L\theta^{-1} \times 1"] \\
L(B \times FA) & L(B \times FA)^A \times A \arrow[l, "\mathrm{ev}"'] & LF(L(B \times FA)^A) \times A \arrow[l, "\epsilon \times 1"']
\end{tikzcd}$$

where λ is the unit of the exponential adjunction, η is the unit of $(L \dashv F)$, and θ is the (invertible) exponential comparison map for F. The verification that this is indeed inverse to $(L\pi_1, \epsilon_A L\pi_2)$ is tedious but straightforward.

Conversely, suppose the Frobenius reciprocity law is satisfied; let $\delta_{A,B}$ denote the inverse of $(L\pi_1, \epsilon_A L\pi_2)$. Since F has a left adjoint, we know that it preserves finite products, so we need only construct an inverse for the exponential

comparison map $\theta\colon F(A^C) \to FA^{FC}$. This time we form the composite

$$\begin{tikzcd}
FA^{FC} \arrow[r, "\eta"] & FL(FA^{FC}) \arrow[r, "F\lambda"] & F((L(FA^{FC}) \times C)^C) \arrow[d, "F(\delta^C)"] \\
F(A^C) & F(LFA^C) \arrow[l, "F(\epsilon^C)"'] & F((L(FA^{FC} \times FC))^C) \arrow[l, "F((L(\mathrm{ev}))^C)"'] .
\end{tikzcd}$$

Again, it is straightforward to verify that this is inverse to θ. □

Two particular cases of 1.5.8 are worth noting:

Corollary 1.5.9 *Let F and L be as in* 1.5.8.

(i) *If F is cartesian closed and L preserves* 1, *then F is full and faithful.*

(ii) *If F is full and faithful and L preserves binary products, then F is cartesian closed.*

Proof (i) Putting $B = 1$ in the Frobenius reciprocity law, we deduce that $\epsilon\colon LF \to 1_C$ is an isomorphism; so F is full and faithful.

(ii) Conversely, if F is full and faithful, then ϵ is an isomorphism, and so the Frobenius reciprocity law follows from the assertion that L preserves binary products. □

In a cartesian closed category $\mathcal{C}$, a class $\mathcal{I}$ of objects (or the full subcategory which it determines) is called an *exponential ideal* if $B \in \mathcal{I}$ implies $B^A \in \mathcal{I}$ for every object A of $\mathcal{C}$.

Lemma 1.5.10 *In any cartesian closed category, the class of subterminal objects (i.e. those B such that $B \to 1$ is monic) is an exponential ideal.*

Proof If B is subterminal, then there is at most one morphism $C \times A \to B$ for any A and C, and hence at most one morphism $C \to B^A$. □

In particular, it follows from 1.5.10 that Sub(1) is a cartesian closed preorder in any cartesian closed category $\mathcal{C}$ (and if $\mathcal{C}$ is locally cartesian closed, then Sub(A) is cartesian closed for all A). In a cartesian closed preorder the exponentiation is normally denoted $(A \Rightarrow B)$ instead of B^A (and called *implication* rather than exponentiation); its defining property is

$$C \leq (A \Rightarrow B) \quad \text{iff} \quad (C \cap A) \leq B \,.$$

A cartesian closed preorder is necessarily locally cartesian closed; this may be seen using 1.5.4, since if $\mathcal{C}$ is a preorder then every object of $\mathcal{C}/B$ is isomorphic to one in the image of B^* (if $A \leq B$, we have $A \cong A \cap B$). Moreover, if $\mathcal{C}$ is a preorder then we have $\mathrm{Sub}_{\mathcal{C}}(A) = \mathcal{C}/A$ for every A, since every morphism of $\mathcal{C}$ is

monic, and so we see that (provided $\mathcal{C}$ is also cocartesian) it is (locally) cartesian closed iff it is a Heyting category as defined after 1.4.10.

A poset (that is, a small preorder in which the only isomorphisms are identities) which is cartesian closed is called a *Heyting semilattice* (sometimes a Brouwerian semilattice); if it is also cocartesian, then it is called a *Heyting algebra*. For future reference, we need to note that the theory of Heyting semilattices is algebraic:

Lemma 1.5.11 *The category* **HSLat** *of Heyting semilattices (and cartesian closed functors) is isomorphic to the category of models of the algebraic theory defined by a constant* 1 *and two binary operations* $\wedge$ *and* $\Rightarrow$, *satisfying the equations*

$$\begin{gathered}1 \wedge x = x, \quad x \wedge x = x, \quad x \wedge y = y \wedge x, \\ x \wedge (y \wedge z) = (x \wedge y) \wedge z, \quad (x \Rightarrow x) = 1, \\ x \wedge (x \Rightarrow y) = x \wedge y, \quad y \wedge (x \Rightarrow y) = y, \\ x \Rightarrow (y \wedge z) = (x \Rightarrow y) \wedge (x \Rightarrow z) .\end{gathered}$$

Proof Let A be a model of the above theory. It is well known that the first four equations are equivalent to saying that A is a meet-semilattice (that is, a cartesian poset) with respect to the partial order defined by $a \leq b$ iff $a = a \wedge b$. So we have to show that the last four equations imply that $a \Rightarrow (-) \colon A \to A$ is right adjoint to $(-) \wedge a$.

First we note that the last equation implies that $a \Rightarrow (-)$ is order-preserving. Now if $c \wedge a \leq b$, we have

$$c = c \wedge (a \Rightarrow c) \leq (a \Rightarrow c) = (a \Rightarrow c) \wedge (a \Rightarrow a) = (a \Rightarrow (c \wedge a)) \leq (a \Rightarrow b),$$

and conversely if $c \leq (a \Rightarrow b)$, then

$$c \wedge a \leq (a \Rightarrow b) \wedge a = a \wedge b \leq b .$$

So A is a Heyting semilattice. Conversely, it is an easy exercise to verify that the last four equations hold in any Heyting semilattice: for example, the last one holds because $a \Rightarrow (-)$, being a right adjoint, preserves limits. □

We may, of course, give a similar algebraic presentation of the theory of Heyting algebras, by adding a further constant 0 and a binary operation $\vee$, plus appropriate equations relating $\vee$ to $\wedge$. Complete Heyting algebras (i.e. ones in which every subset has a least upper bound) are also called *frames* or *locales*; they will be studied extensively in Chapter C1.

We have not yet discussed the relationship between cartesian closedness and the categorical properties studied in Sections A1.3 and A1.4. 'Mere' cartesian closedness does not imply very much: the example of **Cat** shows that it need

not imply regularity. The only positive result we have for merely cartesian closed categories is

Lemma 1.5.12 *If a cartesian closed category $\mathcal{C}$ has an initial object, then it is strict (cf. 1.4.1).*

Proof Write 0 for the initial object, and suppose there exists a morphism $f: A \to 0$. Then the projection $A \times 0 \to A$ is split epic (the splitting being induced by $(1_A, f)$); but $A \times 0 \cong 0$ since $A \times (-)$ preserves colimits. So A is a retract of the initial object, and hence must itself be initial. □

On the other hand, locally cartesian closed categories tend to have most of the good properties of the categories we have studied in previous sections.

Lemma 1.5.13 *Let $\mathcal{C}$ be a locally cartesian closed category.*

(i) *If $\mathcal{C}$ has coequalizers (of effective equivalence relations), then it is regular.*

(ii) *If $\mathcal{C}$ is cocartesian, then it is a Heyting category.*

Proof (i) Since pullback functors in $\mathcal{C}$ have right adjoints, they preserve any coequalizers which exist; so this follows from 1.3.5.

(ii) Let $f: A \to B$ be a morphism of $\mathcal{C}$. Since $\Pi_f: \mathcal{C}/A \to \mathcal{C}/B$ is a right adjoint, it preserves monomorphisms, and hence restricts to a functor $\mathrm{Sub}(A) \to \mathrm{Sub}(B)$, which is clearly right adjoint to $f^*: \mathrm{Sub}(B) \to \mathrm{Sub}(A)$. Moreover, f^* preserves finite unions of subobjects because it has a right adjoint (and the unions exist because $\mathcal{C}$ has finite coproducts; cf. the remarks at the beginning of Section A1.4). □

We cannot hope to have any sort of elementary converse to 1.5.13. Let $\mathbf{Set}_c$ be the full subcategory of **Set** whose objects are all finite or countable sets; then $\mathbf{Set}_c$ is a Boolean pretopos (being closed under the appropriate operations on **Set**), but fails to be cartesian closed for obvious cardinality reasons: there are uncountably many morphisms $1 \times \mathbb{N} \to 2$, but only countably many morphisms $1 \to A$ for any A. Nevertheless, there is a non-elementary converse of sorts: if $\mathcal{C}$ is an ∞-pretopos with a separating set of objects, then it is (a topos, and hence) locally cartesian closed. We shall prove this result in C2.2.8; cf. also 3.4.11.

A category satisfying the hypotheses of 1.5.13 need not be effective as a regular category: for a striking counterexample, see C4.2.4. However, it comes surprisingly close to being positive as a coherent category. We shall say that an object Q of an arbitrary category $\mathcal{C}$ is *quasi-initial* if it admits at most one morphism to any other object of $\mathcal{C}$ (equivalently, if $\mathcal{C}$ has an initial object 0, the unique morphism $0 \to Q$ is epic). And we say that a coproduct $A + B$ is *quasi-disjoint* if the coprojections $A \to A + B$ and $B \to A + B$ are monic, and their intersection in $\mathrm{Sub}(A + B)$ has a quasi-initial object for its domain.

Lemma 1.5.14 *In a locally cartesian closed, cocartesian category $\mathcal{C}$, coproducts are quasi-disjoint. In particular, if 0 has no proper epimorphic images in $\mathcal{C}$ (for example if $\mathcal{C}$ is balanced), then $\mathcal{C}$ is a positive coherent category.*

Proof Given a coproduct $A + B$, form the pullbacks

$$\begin{tikzcd} R \arrow[r, "r"] \arrow[d] & A \arrow[d, "\nu_1"] & S \arrow[l, "s"'] \arrow[d] \\ A \arrow[r, "\nu_1"] & A+B & B \arrow[l, "\nu_2"'] \end{tikzcd} \quad ;$$

then (r, s) is also a coproduct diagram. But r is split epic, being split by the diagonal $\Delta\colon A \rightarrowtail R$; and it is split monic, because the pair of maps $(1_R, \Delta s)$ must factor through (r, s). So it is an isomorphism, and hence ν_1 is monic. Similarly, ν_2 is monic. Now suppose $f_1, f_2\colon S \rightrightarrows C$ are any two morphisms. Then the pairs $(\nu_1 r\colon R \to A + C, \nu_2 f_i\colon S \to A + C)$ $(i = 1, 2)$ have only one possible factorization through (r, s), namely ν_1, since r is epic. So $\nu_2 f_1 = \nu_1 s = \nu_2 f_2$; but we have already seen that ν_2 must be monic, so $f_1 = f_2$. Hence S is quasi-initial.

This proves the first assertion of the lemma; the second follows immediately from it and 1.5.13(ii). The fact that a balanced category $\mathcal{C}$ is a special case of the second assertion follows from the fact that the initial object of $\mathcal{C}$ is strict by 1.5.12, and so any morphism with domain 0 is monic. □

The example of a Heyting algebra, considered as a category, shows that we cannot hope to improve on the result of 1.5.14.

Finally in this section, we return to the regularization construction considered at the end of Section A1.3.

Proposition 1.5.15 *Let $\mathcal{C}$ be a cartesian category. If $\mathcal{C}$ is cartesian closed (resp. locally cartesian closed), then so is* $\mathbf{Reg}(\mathcal{C})$.

Proof First consider the case of 'mere' cartesian closedness. Given two objects $(f_1\colon A_1 \to B_1)$ and $(f_2\colon A_2 \to B_2)$ of $\mathbf{Reg}(\mathcal{C})$, we define the exponential ${f_2}^{f_1}$ to be $(g\colon E \to {B_2}^{A_1})$, where $E \rightarrowtail {A_2}^{A_1}$ is the equalizer of

$$\begin{tikzcd} {A_2}^{A_1} \arrow[r, shift left, "{f_2}^a"] \arrow[r, shift right, "{f_2}^b"'] & {B_2}^{K_1} \end{tikzcd}$$

(and $a, b\colon K_1 \rightrightarrows A_1$ is the kernel-pair of f_1), and g is the composite

$$E \rightarrowtail {A_2}^{A_1} \xrightarrow{{f_2}^{A_1}} {B_2}^{A_1} .$$

The evaluation map $[e]\colon {f_2}^{f_1} \times f_1 \to f_2$ is represented by the composite

$$E \times A_1 \rightarrowtail {A_2}^{A_1} \times A_1 \xrightarrow{\mathrm{ev}} A_2;$$

we must check that this does indeed represent a morphism of $\mathbf{Reg}(\mathcal{C})$, i.e. that its composite with f_2 coequalizes $(c \times a)$ and $(d \times b)$ where $c, d\colon K_2 \rightrightarrows E$ is the

kernel-pair of g. But this composite can also be written as

$$E \times A_1 \xrightarrow{g \times 1} B_2{}^{A_1} \times A_1 \xrightarrow{\mathrm{ev}} B_2,$$

so it coequalizes $(c \times 1)$ and $(d \times 1)$; and it coequalizes $(1 \times a)$ and $(1 \times b)$ by the definition of E.

Now, given $[h]\colon f_3 \times f_1 \to f_2$ in $\mathbf{Reg}(\mathcal{C})$, let $\tilde{h}\colon A_3 \to E$ be the factorization of the transpose $\overline{h}\colon A_3 \to A_2{}^{A_1}$ through the equalizer (which exists because $f_2 h$ has equal composites with $(1 \times a)$ and $(1 \times b)$, by assumption). Then $\tilde{h}$ defines a morphism $f_3 \to g$ in $\mathbf{Reg}(\mathcal{C})$, since $(f_2{}^{A_1})h$ coequalizes the kernel-pair of f_3. We need to check that the equivalence class $[\tilde{h}]$ is independent of the choice of h; but if $[h_1] = [h_2]$ (that is, $f_2 h_1 = f_2 h_2$), then clearly $f_2{}^{A_1}$ coequalizes $\overline{h_1}$ and $\overline{h_2}$, and so g coequalizes $\widetilde{h_1}$ and $\widetilde{h_2}$. Also, we have $[e]([\tilde{h}] \times 1) = [\mathrm{ev}(\overline{h} \times 1)] = [h]$; and if $k\colon A_3 \to E$ is any morphism such that $[e]([k] \times 1) = [h]$, then $\mathrm{ev}(k \times 1)$ and h represent the same morphism $f_3 \times f_1 \to f_2$, whence k and $\tilde{h}$ have equal composites with g. So we have verified that $f_2{}^{f_1}$ has the universal property of an exponential in $\mathbf{Reg}(\mathcal{C})$.

The locally cartesian closed case is similar but more complicated. We shall content ourselves with giving the construction of the exponential $[g_2]^{[g_1]}$ (where $[g_1]\colon f_1 \to f_3$ and $[g_2]\colon f_2 \to f_3$) in $\mathbf{Reg}(\mathcal{C})/f_3$, leaving the reader to verify that it has the appropriate universal property. Let $a, b\colon K_1 \rightrightarrows A_1$ be the kernel-pair of $f_1\colon A_1 \to B_1$, and let k denote the composite $f_3 g_1 a = f_3 g_1 b$. Now form the equalizer

$$E \rightarrowtail (f_3 g_2)^{(f_3 g_1)} \underset{(\overline{f_2})^b}{\overset{(\overline{f_2})^a}{\rightrightarrows}} (B_3^*(B_2))^k$$

in $\mathcal{C}/B_3$, where $\overline{f_2}\colon f_3 g_2 \to B_3^*(B_2)$ is the transpose of $f_2\colon A_2 = \Sigma_{B_3}(f_3 g_2) \to B_2$; and let $g_4\colon A_4 \to A_3$ denote the pullback $f_3^*(E)$ in $\mathcal{C}/A_3$. Define B_4 to be $\Sigma_{B_3}((B_3^*(B_2))^{(f_3 g_1)})$, and let $f_4\colon A_4 \to B_4$ be the map obtained by applying Σ_{B_3} to the composite

$$\Sigma_{f_3} f_3^*(E) \longrightarrow E \rightarrowtail (f_3 g_2)^{(f_3 g_1)} \xrightarrow{(\overline{f_2})^{(f_3 g_1)}} (B_3^*(B_2))^{(f_3 g_1)}$$

in $\mathcal{C}/B_3$, where the first factor is the counit of $(\Sigma_{f_3} \dashv f_3^*)$. Then we claim that $g_4\colon A_4 \to A_3$ defines a morphism $[g_4]\colon f_4 \to f_3$ in $\mathbf{Reg}(\mathcal{C})$, and that $[g_4]$ has the universal property of an exponential $[g_2]^{[g_1]}$ in $\mathbf{Reg}(\mathcal{C})/f_3$. The remaining details are tedious but routine verification. □

Suggestions for further reading: Carboni & Rosolini [230], Johnstone [531], Lambek [668], Lambek & Scott [682], Rosický [1060].

A1.6 Subobject classifiers

In the earlier sections of this chapter, we have been much concerned with properties of subobjects in various classes of categories. Many problems concerning subobjects are simplified if we assume that they are all obtained by pullback from one particular one. In this section we investigate some of the consequences of that assumption.

Let $\mathcal{C}$ be a category with pullbacks. A *generic subobject* in $\mathcal{C}$ is a monomorphism $\top: B' \rightarrowtail B$ such that, given any monomorphism $m: A' \rightarrowtail A$ in $\mathcal{C}$, there is a unique $f: A \to B$ for which we have a pullback square

$$\begin{tikzcd} A' \arrow[r] \arrow[d, tail, "m"] & B' \arrow[d, tail, "\top"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

(i.e. such that $f^*(\top) \cong m$ in $\mathrm{Sub}(A)$). The symbol $\top$ is pronounced 'true' or 'top', according to taste; the reason for this nomenclature will become clear later. f is called the *classifying map* or *characteristic morphism* of the subobject m; it is clear that two subobjects of A have the same classifying map iff they are isomorphic in $\mathrm{Sub}(A)$, and hence (recalling our convention that a category with pullbacks comes equipped with a choice of 'canonical pullbacks') a category with a generic subobject comes equipped with a choice of canonical representatives for the isomorphism classes of objects of $\mathrm{Sub}(A)$, for any A. (Thus, in such a category, it does even less harm than usual to confuse individual monomorphisms with isomorphism classes thereof.)

Lemma 1.6.1 *The domain of a generic subobject is a terminal object of $\mathcal{C}$.*

Proof For any A, a morphism $f: A \to B$ factors through $\top: B' \rightarrowtail B$ iff $f^*(\top)$ is an isomorphism; so there is exactly one morphism which does so. □

In particular, a category with a generic subobject necessarily has a terminal object; so there was spurious generality in our initial assumption that $\mathcal{C}$ had pullbacks rather than all finite limits. (It is possible to gain still further spurious generality by requiring $\mathcal{C}$ merely to have binary products, plus arbitrary pullbacks of the particular monomorphism $\top$: for the equalizer of a parallel pair $f, g: A \rightrightarrows A'$ may be obtained as the pullback of $\top$ along $\delta(f,g)$, where $\delta: A' \times A' \to B$ is the classifying map of the diagonal subobject $A' \rightarrowtail A' \times A'$.) The codomain of a generic subobject is traditionally denoted Ω, and called a *subobject classifier* for $\mathcal{C}$; from now on, therefore, we shall write the generic subobject as $\top: 1 \rightarrowtail \Omega$. One consequence of 1.6.1 is worth remarking immediately:

Corollary 1.6.2 *In a category with a subobject classifier, every monomorphism is regular. In particular, such a category is balanced.*

Proof The generic subobject is regular monic, because it is split by the unique morphism $\Omega \to 1$. But in any category, a pullback of a regular monomorphism is regular monic: if m is an equalizer of g and h, then $f^*(m)$ is an equalizer of gf and hf. The second assertion follows from the first as in the proof of 1.4.9. □

Lemma 1.6.3 *Let C be a cartesian category with a subobject classifier Ω.*

(i) *For every object A of C, the preorder* Sub(A) *is cartesian closed; and in fact Ω has the structure of an internal Heyting semilattice in C (cf.* 1.5.11*).*

(ii) *If C has finite unions of subobjects which are stable under pullback (for example, if C is a coherent category), then Ω is an internal Heyting algebra in C.*

Proof (i) Let $\wedge : \Omega \times \Omega \to \Omega$ be the classifying map of the subobject $(\top, \top) : 1 \rightarrowtail \Omega \times \Omega$. We note that this subobject is the intersection of $\pi_1^*(\top)$ and $\pi_2^*(\top)$, where π_1 and π_2 are the product projections $\Omega \times \Omega \rightrightarrows \Omega$; and since intersections of subobjects are stable under pullback, we see that if $f, f' : A \rightrightarrows \Omega$ are the classifying maps of subobjects $A' \rightarrowtail A$ and $A'' \rightarrowtail A$ respectively, then the composite $\wedge(f, f')$ classifies $A' \cap A'' \rightarrowtail A$. Since the composite $\top A : A \to 1 \to \Omega$ classifies the terminal object of Sub(A), it is now easy to verify that the constant $\top$ and the binary operation $\wedge$ give Ω the structure of an internal meet-semilattice (i.e. they satisfy the diagrammatic forms of the first four equations in the statement of 1.5.11). Now let $\Omega_1 \rightarrowtail \Omega \times \Omega$ be the equalizer of $\wedge$ and π_1, and let $\Rightarrow \Omega \times \Omega \to \Omega$ be the classifying map of this subobject. Then we see that if f and f' are the classifying maps of subobjects $A' \rightarrowtail A$ and $A'' \rightarrowtail A$, the morphism $(f, f') : A \to \Omega \times \Omega$ factors through $\Omega_1 \rightarrowtail \Omega \times \Omega$ iff $A' \cap A'' \cong A'$, i.e. iff $A' \leq A''$ in Sub(A). If now $m : B \rightarrowtail A$ is a third subobject of A, classified by $g : A \to \Omega$, and $(A' \Rightarrow A'') \rightarrowtail A$ denotes the subobject classified by the composite $\Rightarrow (f, f')$, we see that

$$
\begin{aligned}
B \leq (A' \Rightarrow A'') \text{ in } \mathrm{Sub}(A) \quad &\text{iff} \quad \Rightarrow (f, f')m = \top B \\
&\text{iff} \quad (f, f')m \text{ factors through } \Omega_1 \rightarrowtail \Omega \times \Omega \\
&\text{iff} \quad m^*(A') \leq m^*(A'') \text{ in } \mathrm{Sub}(B) \\
&\text{iff} \quad A' \cap B = \exists_m m^*(A') \leq A'' \text{ in } \mathrm{Sub}(A)\,.
\end{aligned}
$$

So we have verified that Sub(A) is cartesian closed for any A; and since its implication operation is induced by the binary operation $\Rightarrow$ on Ω, it is again straightforward to verify that this operation, with $\top$ and $\wedge$, satisfies the last four equations of 1.5.11.

(ii) Let $\bot : 1 \to \Omega$ (pronounced 'false' or 'bottom') classify the least subobject $0 \rightarrowtail 1$ of 1, and let $\vee : \Omega \times \Omega \to \Omega$ classify the union of the subobjects $\pi_1^*(\top)$ and $\pi_2^*(\top)$. Since unions in C are stable under pullback, they are induced by these operations on Ω in the same sense that intersections are induced by $\top$ and $\wedge$; so

it is again easy to verify that Ω is an internal lattice in $\mathcal{C}$, and hence an internal Heyting algebra. $\square$

In the presence of a subobject classifier, the verification that a category is a Heyting category becomes appreciably simplified:

Proposition 1.6.4 *Let $\mathcal{C}$ be a regular category having a subobject classifier Ω. Then $\mathcal{C}$ is a Heyting category iff it 'admits universal quantification over Ω' in the sense that, for every object A, the pullback functor $\pi_1^*\colon \mathrm{Sub}(A) \to \mathrm{Sub}(A \times \Omega)$ has a right adjoint $\forall_{\pi_1}$. In particular, if Ω is exponentiable in $\mathcal{C}$, then $\mathcal{C}$ is a Heyting category.*

Proof 'Only if' is obvious, so we need to prove 'if'. First we define $0 \rightarrowtail 1$ to be $\forall_\Omega(\top\colon 1 \rightarrowtail \Omega)$ (taking $A = 1$ in the condition above). Then the counit $\Omega^*(0) = \Omega^*\forall_\Omega(\top) \leq \top$ ensures that, for any subobject $m\colon A' \rightarrowtail A$ in $\mathcal{C}$, we have $A^*(0) = \chi_m^*\Omega^*(0) \leq \chi_m^*(\top) = m$ in $\mathrm{Sub}(A)$, where χ_m is the classifying map of m. So the posets $\mathrm{Sub}(A)$ have least elements, which are stable under pullback.

To obtain binary unions, we first form the subobject of $\Omega \times \Omega \times \Omega$ classified by the map $v = ((\pi_1 \Rightarrow \pi_3) \Rightarrow ((\pi_2 \Rightarrow \pi_3) \Rightarrow \pi_3))$, where $\Rightarrow$ is the Heyting implication constructed in 1.6.3(i), and then apply universal quantification along $\pi_{12}\colon \Omega \times \Omega \times \Omega \to \Omega \times \Omega$. Let $\vee\colon \Omega \times \Omega \to \Omega$ classify the resulting subobject. An easy calculation shows that we have $\pi_1 \leq v$ and $\pi_2 \leq v$ in the poset $\mathcal{C}(\Omega^3, \Omega) \cong \mathrm{Sub}(\Omega^3)$, and hence $\pi_1 \leq \vee$ and $\pi_2 \leq \vee$ in $\mathcal{C}(\Omega^2, \Omega)$. Thus, given any two subobjects $m_i\colon A_i \rightarrowtail A$ $(i = 1, 2)$ in $\mathcal{C}$, the subobject classified by the composite $\vee(\chi_{m_1}, \chi_{m_2})$ is an upper bound for m_1 and m_2 in $\mathrm{Sub}(A)$. But if $m_3\colon A_3 \rightarrowtail A$ is any such upper bound, then the composite

$$A \xrightarrow{(\chi_{m_1}, \chi_{m_2}, \chi_{m_3})} \Omega \times \Omega \times \Omega \xrightarrow{\ v\ } \Omega$$

reduces to χ_{m_3}; that is, $(\chi_{m_1}, \chi_{m_2}, \chi_{m_3})$ factors through the equalizer of v and π_3. Hence the inequality $\vee\pi_{12} \leq v$, which is the counit of the adjunction $(\pi_{12}^* \dashv \forall_{\pi_{12}})$ translated into terms of classifying maps, yields $\vee(\chi_{m_1}, \chi_{m_2}) \leq \chi_{m_3}$ in $\mathcal{C}(A, \Omega)$. So the posets $\mathrm{Sub}(A)$ have binary joins, which are stable under pullback; thus $\mathcal{C}$ is a coherent category.

It remains to construct universal quantification along an arbitrary morphism $f\colon A \to B$ of $\mathcal{C}$. Given a subobject $m\colon A' \rightarrowtail A$, we first form a subobject $U \rightarrowtail A \times B \times \Omega$ by pulling back the graphs $(1, f)\colon A \rightarrowtail A \times B$ and $(1, \chi_m)\colon A \rightarrowtail A \times \Omega$ along the appropriate product projections and forming their intersection. Then we apply existential quantification along the projection $\pi_{23}\colon A \times B \times \Omega \to B \times \Omega$ to obtain $V \rightarrowtail B \times \Omega$, and let $V' \rightarrowtail B \times \Omega$ be the subobject classified by $(v \Rightarrow \pi_2)$, where v is the classifying map of V. Finally, we apply universal quantification along $\pi_1 : B \times \Omega \to B$ to V', to obtain a subobject $W \rightarrowtail B$; we claim that this subobject is the universal quantification $\forall_f(A')$.

First we show that $f^*(W) \leq A'$ in $\mathrm{Sub}(A)$. For this, consider the diagram

$$\begin{tikzcd}
\mathrm{Sub}(A \times B \times \Omega) \arrow[r, "\exists_{\pi_{23}}"] \arrow[d, "(1\times f\times 1)^*"] & \mathrm{Sub}(B \times \Omega) \arrow[r, "((-)\Rightarrow B)"] \arrow[d, "(f\times 1)^*"] & \mathrm{Sub}(B \times \Omega) \arrow[r, "\forall_{\pi_1}"] \arrow[d, "(f\times 1)^*"'] & \mathrm{Sub}(B) \arrow[d, "f^*"] \\
\mathrm{Sub}(A \times A \times \Omega) \arrow[r, "\exists_{\pi_{23}}"] & \mathrm{Sub}(A \times \Omega) \arrow[r, "((-)\Rightarrow A)"] & \mathrm{Sub}(A \times \Omega) \arrow[r, "\forall_{\pi_1}"] & \mathrm{Sub}(A)
\end{tikzcd}$$

where the outer squares commute by the Beck–Chevalley conditions (note that the Beck–Chevalley condition for $\forall_{\pi_1}$ follows from that for $\exists_f$, as in 1.4.11), and the central square commutes by the stability of implication under pullback (here A and B denote the subobjects $(1, \top A)\colon A \rightarrowtail A \times \Omega$ and $(1, \top B)\colon B \rightarrowtail B \times \Omega$ classified by the second projections). If we chase the subobject U around the top and right-hand edges of the diagram, we obtain $f^*(W)$; but if we chase it down the left-hand edge, we obtain a subobject which contains the subobject $(1, 1, \chi_m)\colon A \rightarrowtail A \times A \times \Omega$, and whose image under π_{23} therefore contains $(1, \chi_m)\colon A \rightarrowtail A \times \Omega$. Since implication is order-reversing in its first variable, the rest of the bottom edge sends this to a subobject contained in $\forall_{\pi_1}((1, \chi_m) \Rightarrow (1, \top A))$, that is the largest subobject $A'' \rightarrowtail A$ such that $(A'' \times \Omega) \cap (1, \chi_m) \leq (1, \top A)$. But the latter is easily seen to be the equalizer of χ_m and $\top A$ – that is, $A' \rightarrowtail A$.

Now suppose we have a subobject $B' \rightarrowtail B$ such that $f^*(B') \leq A'$ in $\mathrm{Sub}(A)$. To show $B' \leq W$ in $\mathrm{Sub}(B)$, we need to show that $B' \times \Omega \leq V'$ in $\mathrm{Sub}(B \times \Omega)$, or equivalently that $(B' \times \Omega) \cap V \leq (1, \top B)$ in $\mathrm{Sub}(B \times \Omega)$. By the Frobenius reciprocity condition 1.3.3, this is in turn equivalent to showing that $(A \times B' \times \Omega) \cap U \leq ((1, \top(A \times B))\colon A \times B \rightarrowtail A \times B \times \Omega)$ in $\mathrm{Sub}(A \times B \times \Omega)$. But the condition $f^*(B') \leq A'$ implies that the intersection of $A \times B'$ with the graph of f in $\mathrm{Sub}(A \times B)$ is contained in $A' \times B'$; and the intersecton of $A' \times \Omega$ with the graph of χ_m in $\mathrm{Sub}(A \times \Omega)$ is contained in $(1, \top A)\colon A \rightarrowtail A \times \Omega$. So the inequality we require follows from the definition of U (plus stability of intersections under pullback).

Finally, suppose Ω is exponentiable in $\mathcal{C}$. Then by 1.5.2(ii) $A^*(\Omega)$ is exponentiable in $\mathcal{C}/A$ for all A, whence by 1.5.2(i) we obtain functors $\Pi_{\pi_1} : \mathcal{C}/(A \times \Omega) \to \mathcal{C}/A$ right adjoint to the pullback functors. Restricting these to subobjects, as in 1.5.13(ii), yields the required functors $\forall_{\pi_1}$. □

We have just seen that the existence of the exponentials A^Ω for all A, in a regular category with subobject classifier Ω, has important consequences. In the next chapter, we shall see that the existence of Ω^A for all A has even more far-reaching consequences (cf. 2.2.7 and 2.3.4). Note also that the first part of the proof of 1.6.4 does not make any use of regularity; thus a cartesian category with subobject classifier, which admits universal quantification over Ω, automatically has pullback-stable finite unions of subobjects.

So far we have not seen any examples of subobject classifiers. However, we have

Lemma 1.6.5 *A Boolean coherent category in which the coproduct* $1+1$ *exists and is disjoint (in particular, a Boolean pretopos) has a subobject classifier.*

Proof Define $\Omega = 1+1$ and let $\top \colon 1 \rightarrowtail \Omega$ be the first coprojection (and write $\perp$ for the other coprojection). Given a subobject $A' \rightarrowtail A$, let $A'' \rightarrowtail A$ be its complement; then A is the coproduct of A' and A'' by 1.4.4, so there is a unique $f \colon A \to \Omega$ such that the diagram

$$\begin{tikzcd} A' \arrow[r, tail] \arrow[d] & A \arrow[d, "f"] & A'' \arrow[l, tail] \arrow[d] \\ 1 \arrow[r, tail, "\top"] & \Omega & 1 \arrow[l, tail, "\perp"'] \end{tikzcd}$$

commutes. Now $f^*(\top)$ is a subobject of A containing A' and disjoint from A'', so it must be (isomorphic to) A'; but conversely any morphism $f \colon A \to \Omega$ which makes the left-hand square above a pullback must make the right-hand square commute, since $\exists_f(A'')$ must be contained in the complement of $\top$. So there is a unique such f. □

In particular, **Set** has a subobject classifier, as does $[\mathcal{C}, \mathbf{Set}]$ for any groupoid $\mathcal{C}$ (not necessarily locally small; note that we did not make use of local smallness of $\mathcal{C}$ in the first half of the proof of 1.4.12). But we also have

Lemma 1.6.6 *For any small category* $\mathcal{C}$, *the functor category* $[\mathcal{C}, \mathbf{Set}]$ *has a subobject classifier.*

Proof As in the proof of 1.5.5, the Yoneda Lemma determines what Ω must be: elements of $\Omega(A)$ correspond to morphisms $\mathcal{C}(A,-) \to \Omega$ in $[\mathcal{C}, \mathbf{Set}]$, and hence to (isomorphism classes of) subobjects of $\mathcal{C}(A,-)$. Now a subfunctor F of $\mathcal{C}(A,-)$ is determined by the set

$$R = \{f \in \text{mor}\,\mathcal{C} \mid \text{dom}\,f = A \text{ and } f \in F(\text{cod}\,f)\};$$

R is a *cosieve* on A, i.e. a set of morphisms with common domain A such that $f \in R$ implies $gf \in R$ for any g composable with f, and any cosieve on A may arise in this way. So we define $\Omega(A)$ to be the set of cosieves on A; given a morphism $f \colon A \to B$ and $R \in \Omega(A)$, we define

$$\Omega(f)(R) = \{g \in \text{mor}\,\mathcal{C} \mid \text{dom}\,g = B \text{ and } gf \in R\}.$$

It is clear that this makes Ω into a functor $\mathcal{C} \to \mathbf{Set}$. For any A, we define $\top_A$ to be the cosieve of all morphisms with domain A; then $\top$ is a natural

transformation from 1 (i.e. the constant functor with value 1) to Ω. Now, given a subobject $F' \rightarrowtail F$ in $[\mathcal{C}, \mathbf{Set}]$, we define $\phi \colon F \to \Omega$ by

$$\phi_A(x) = \{f \mid \mathrm{dom}\ f = A \text{ and } F(f)(x) \in F'(\mathrm{cod}\ f)\}$$

for $x \in F(A)$; it is clear that ϕ is a natural transformation $F \to \Omega$, and that $\phi_A(x) = \top_A$ iff $x \in F'(A)$, i.e. that

$$\begin{tikzcd}
F' \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\
F \arrow[r, "\phi"] & \Omega
\end{tikzcd}$$

is a pullback. But ϕ is the unique natural transformation with this property; for, given a cosieve R, we have $f \in R$ iff $\Omega(f)(R) = \top_{\mathrm{cod}\ f}$, and hence if ψ is a natural transformation making the square a pullback we have $f \in \psi_{\mathrm{dom}\ f}(x)$ iff $\psi_{\mathrm{cod}\ f}(F(f)(x)) = \top_{\mathrm{cod}\ f}$, iff $F(f)(x) \in F'(\mathrm{cod}\ f)$. So Ω is a subobject classifier for $[\mathcal{C}, \mathbf{Set}]$. □

The category $\mathbf{LH}/X$ also has a subobject classifier for any X; once again, we defer the proof of this until 4.2.4(*e*). Possession of a subobject classifier is stable under slicing, in an obvious sense:

Lemma 1.6.7 *If Ω is a subobject classifier in $\mathcal{C}$, then $B^*(\Omega)$ is a subobject classifier in $\mathcal{C}/B$, for any object B of $\mathcal{C}$.*

Proof Let $f \colon A \to B$ be an object of $\mathcal{C}/B$, and

$$\begin{tikzcd}
A' \arrow[r, tail, "m"] \arrow[dr, "f'"'] & A \arrow[d, "f"] \\
& B
\end{tikzcd}$$

a subobject of f (note that the forgetful functor $\mathcal{C}/B \to \mathcal{C}$ preserves and reflects monomorphisms). Let $g \colon A \to \Omega$ be the classifying map of $m \colon A' \rightarrowtail A$ in $\mathcal{C}$; then $(g, f) \colon A \to \Omega \times B$ defines a morphism $f \to B^*(\Omega)$ in $\mathcal{C}/B$, which is easily seen (since $\mathcal{C}/B \to \mathcal{C}$ preserves and reflects pullbacks) to be the unique morphism

making

$$\begin{tikzcd} f' \arrow[r] \arrow[d, tail, "m"] & B^*(1) \arrow[d, tail, "B^*(\top)"] \\ f \arrow[r] & B^*(\Omega) \end{tikzcd}$$

a pullback. □

Rather remarkably, it is also stable under co-slicing (unlike many of the properties we have considered up to now); we shall not need this result subsequently, but we include it for completeness.

Lemma 1.6.8 *If Ω is a subobject classifier in $\mathcal{C}$, then $\top B\colon B \to 1 \to \Omega$ is a subobject classifier in the co-slice category $B\backslash\mathcal{C}$.*

Proof Let

$$\begin{tikzcd} B \arrow[d, "f'"'] \arrow[dr, "f"] & \\ A' \arrow[r, tail, "m"] & A \end{tikzcd}$$

be a monomorphism in $B\backslash\mathcal{C}$, and let $g\colon A \to \Omega$ be the classifying map of $m\colon A' \rightarrowtail A$ in $\mathcal{C}$. Since f factors through m, the composite gf is equal to $\top B$, i.e. g is a morphism $f \to \top B$ in $B\backslash\mathcal{C}$. The rest is straightforward. □

The following result, though seemingly rather special, will be of use in various contexts later on.

Lemma 1.6.9 *Let $\mathcal{C}$ and $\mathcal{D}$ be cartesian categories such that $\mathcal{C}$ has a subobject classifier; let $S, T\colon \mathcal{C} \rightrightarrows \mathcal{D}$ be a pair of cartesian functors, and $\theta\colon S \to T$ a natural transformation. Then the following are equivalent:*

(i) *θ is pointwise monic, i.e. θ_A is monic for each $A \in \mathrm{ob}\,\mathcal{C}$.*

(ii) *For each monomorphism $m\colon A' \rightarrowtail A$ in $\mathcal{C}$, the naturality square*

$$\begin{tikzcd} SA' \arrow[r, "\theta_{A'}"] \arrow[d, tail, "Sm"] & TA' \arrow[d, tail, "Tm"] \\ SA \arrow[r, "\theta_A"] & TA \end{tikzcd}$$

is a pullback.

Moreover, if these conditions hold, then θ_A is an isomorphism for each subterminal object A of $\mathcal{C}$.

Proof (i) $\Rightarrow$ (ii): Consider first the generic monomorphism $\top : 1 \rightarrowtail \Omega$. In the naturality square

$$\begin{tikzcd} S1 \arrow[r, "\theta_1"] \arrow[d, tail, "S\top"] & T1 \arrow[d, tail, "T\top"] \\ S\Omega \arrow[r, "\theta_\Omega"] & T\Omega \end{tikzcd}$$

the top edge is an isomorphism since S and T both preserve 1, and the bottom edge is monic by assumption; it is easy to see that these conditions force it to be a pullback. Now every $m: A' \rightarrowtail A$ is a pullback of $\top$, and S and T both preserve pullbacks; so a straightforward diagram-chase shows that the naturality square for m is also a pullback.

(ii) $\Rightarrow$ (i): Given an object A of $\mathcal{C}$, apply (ii) to the diagonal morphism $\Delta: A \rightarrowtail A \times A$. Since both S and T preserve binary products, this yields a pullback square

$$\begin{tikzcd} SA \arrow[r, "\theta_A"] \arrow[d, tail, "\Delta"] & TA \arrow[d, tail, "\Delta"] \\ SA \times SA \arrow[r, "\theta_A \times \theta_A"] & TA \times TA \end{tikzcd} ;$$

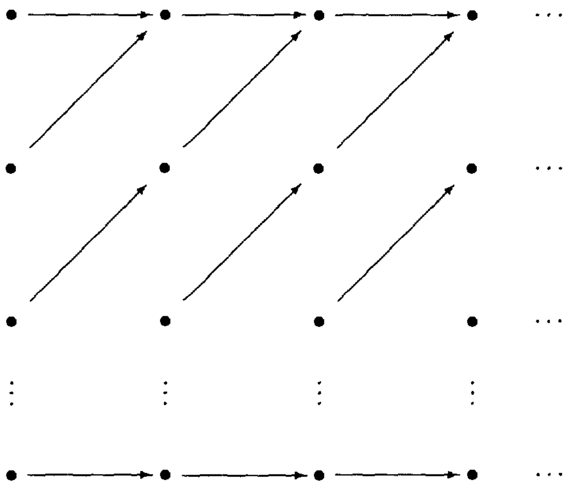
but this says precisely that θ_A is monic.

The final assertion is immediate from (ii), since we have already noted that θ_1 must be an isomorphism, and θ_A is a pullback of θ_1 if A is subterminal. □

In contrast to 1.5.15, the regularization $\mathbf{Reg}(\mathcal{C})$ of a cartesian category $\mathcal{C}$ does not in general inherit a subobject classifier from $\mathcal{C}$. The reason may be seen from 1.3.10(d): if we take $\mathcal{C}$ to be the category of directed graphs, then $\mathcal{C}$ has a subobject classifier by 1.6.6, but $\mathbf{Reg}(\mathcal{C})$ does not because, as we saw after 1.4.17, it is locally small but not well-powered.

The example of a functor category $[\mathcal{C}, \mathbf{Set}]$ shows that a subobject classifier need not be finite in any very strong sense: for example if $\mathcal{C}$ is the ordered set $\mathbf{N}$ of natural numbers, then there are infinitely many cosieves on each number n, namely the empty cosieve and the cosieve of all numbers $\geq m$, for each $m \geq n$.

We can picture the transition maps $\Omega(n) \to \Omega(n+1)$ as follows:



Here the generic subobject is the morphism $1 \to \Omega$ corresponding to the top row of dots; the bottom row corresponds to the empty cosieve.

However, Ω does satisfy Dedekind's definition of finiteness: all its monic endomorphisms are isomorphisms (although it may have epic endomorphisms which are not invertible, as can be seen from the picture above). In fact we can be more specific:

Lemma 1.6.10 *Let $f\colon \Omega \rightarrowtail \Omega$ be a monomorphism. Then $ff = 1_\Omega$.*

Proof First form the pullbacks

$$\begin{tikzcd}
U \arrow[r] \arrow[d, tail, "g"'] & 1 \arrow[d, tail, "\top"] \\
\Omega \arrow[r, "f"] & \Omega
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
V \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\
U \arrow[r, "g"] & \Omega
\end{tikzcd}
\quad .$$

Since f is monic, U is a subobject of 1, and V is a subobject of 1 contained in U. Now in the diagram

$$\begin{tikzcd}
V \arrow[r] \arrow[d, tail] & V \arrow[r] \arrow[d, tail] & U \arrow[r] \arrow[d, tail, "g"] & 1 \arrow[d, tail, "\top"] \\
U \arrow[r] & 1 \arrow[r, "\top"] & \Omega \arrow[r, "f"] & \Omega
\end{tikzcd}$$

all three squares are pullbacks, so the bottom composite must equal g. Thus $ffg = f\top U = g$, i.e. the square

$$\begin{tikzcd} U \arrow[r, "1"] \arrow[d, tail, "g"] & U \arrow[d, tail, "g"] \\ \Omega \arrow[r, "ff"] & \Omega \end{tikzcd}$$

commutes; but since ff is monic it must be a pullback, and hence $fff = f$ since both classify the subobject g of Ω. But f is monic, so $ff = 1_\Omega$. □

Note that the object U in the above proof is in fact isomorphic to 1. In the opposite direction, let $V \rightarrowtail 1$ be a subterminal object, and let $v\colon 1 \to \Omega$ be its classifying map. v itself must be monic, and so has a classifying map $f\colon \Omega \to \Omega$; when is f an isomorphism (equivalently, monic)? It is not hard to verify that f is the composite

$$\Omega \cong \Omega \times 1 \xrightarrow{1 \times v} \Omega \times \Omega \xrightarrow{\Leftrightarrow} \Omega$$

where $\Leftrightarrow$ is the bi-implication of the Heyting semilattice structure on Ω (cf. 1.6.3), or equivalently the classifying map of the diagonal subobject $\Omega \rightarrowtail \Omega \times \Omega$. Now, in any Heyting semilattice H, we have

$$\begin{aligned}((u \Leftrightarrow v) \Leftrightarrow v) &= (((u \Rightarrow v) \wedge (v \Rightarrow u)) \Rightarrow v) \wedge (v \Rightarrow ((u \Rightarrow v) \wedge (v \Rightarrow u))) \\ &= ((v \Rightarrow u) \Rightarrow ((u \Rightarrow v) \Rightarrow v)) \wedge (v \Rightarrow (u \Rightarrow v)) \wedge (v \Rightarrow (v \Rightarrow u)) \\ &= ((v \Rightarrow u) \Rightarrow ((u \Rightarrow v) \Rightarrow v)) \wedge 1 \wedge (v \Rightarrow u) \\ &= ((u \Rightarrow v) \Rightarrow v) \wedge (v \Rightarrow u)\ .\end{aligned}$$

It follows that we always have $((u \Leftrightarrow v) \Leftrightarrow v) \geq u$, and that we have equality here iff $((u \Rightarrow v) \Rightarrow v) \leq ((v \Rightarrow u) \Rightarrow u)$. If this condition holds for all u, we say v is a *widespread* element of H; in the case when H is a Heyting algebra (i.e. has finite joins as well as meets), it is equivalent to saying that the principal filter $\uparrow(v) = \{u \in H \mid v \leq u\}$ is a complemented lattice (and hence a Boolean algebra). From the above discussion, we may conclude

Corollary 1.6.11 *In a cartesian category with a subobject classifier Ω, the automorphisms of Ω correspond bijectively to (isomorphism classes of) subobjects V of* 1 *which are 'internally widespread', in the sense that $V \times A \rightarrowtail A$ is widespread in* Sub(A) *for all A.* □

Further, the group operation on automorphisms corresponds to the bi-implication operation on subobjects of 1; it is again easy to verify that, in any Heyting semilattice, this operation is associative when restricted to the widespread elements, and it makes them into a group of exponent 2.

For example, let $\mathcal{C}$ be a small category. Subobjects of 1 in $[\mathcal{C}, \mathbf{Set}]$ correspond to cosieves on $\mathcal{C}$ (that is, to subsets $R \subseteq \text{ob}\,\mathcal{C}$ such that the codomain of any morphism with domain in R is also in R), and it may be verified that a cosieve R is internally widespread iff every morphism with codomain *not* in R is an isomorphism; equivalently, iff the objects not in R are all strict in the sense defined after 1.4.1. Thus the group of automorphisms of Ω may be identified with the set of isomorphism-closed subsets of the set of all strict objects of $\mathcal{C}$ (the group operation being symmetric difference).

We shall meet the notion of widespread subobject again, in a different context, in Section D4.5.

Suggestions for further reading: Borceux & Veit [168], Engenes [335], Johnstone [506, 531].

A2

TOPOSES – BASIC THEORY

A2.1 Definition and examples

We are now ready to present the definition of an (elementary) topos, which is the central notion of this book: it is a category having all the good 'set-like' properties discussed in the previous chapter, except (possibly) for Booleanness. It is a remarkable fact that all these properties are implied by the conjunction of just two of them:

Definition 2.1.1 A *topos* is a properly cartesian closed category with a subobject classifier.

The proof that this implies all the other good properties will occupy much of Sections A2.2–A2.4; the present section is largely devoted to setting up some examples which we shall use later on in the book. Before that, however, we must mention that the above definition of a topos, though it is appreciably simplified from that originally adopted by F. W. Lawvere and M. Tierney when they began the subject in 1969, is not the simplest possible; one can combine the exponentials and the subobject classifier into a single structure, that of power objects.

Let $\mathcal{E}$ be a topos (as defined above), and A an object of $\mathcal{E}$. We write PA for the exponential Ω^A, and $\in_A \rightarrowtail PA \times A$ for the subobject classified by the evaluation map $\mathrm{ev}\colon PA \times A \to \Omega$. (The reason for this notation is that, in **Set**, PA may be identified with the power-set of A – recall that Ω is just a two-element set in this case – and $\in_A$ is then the subset

$$\{(A', a) \mid a \in A'\},$$

i.e. the membership relation between subsets and elements of A. Admittedly, the latter is 'back to front', which has led some authors to denote it by $\ni_A$ (or ${}_A\ni$) rather than $\in_A$, but we prefer to use the more familiar symbol.) This subobject has the following universal property: given an object B and a subobject $R \rightarrowtail B \times A$ (we think of R as a *relation* from B to A, a point of view which will be extensively explored in Chapter A3), there is a unique morphism $r\colon B \to PA$ (namely the exponential transpose of the classifying map $B \times A \to \Omega$ of R) for

which there is a pullback square

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] & \in_A \arrow[d, tail] \\ B \times A \arrow[r, "r \times 1"] & PA \times A \end{tikzcd}$$

In a general cartesian category $\mathcal{E}$, by a *power object* of an object A we mean an object PA equipped with a subobject $\in_A \rightarrowtail PA \times A$ satisfying the above universal property. (We call the morphism r above the *name* of the relation R; as with subobject classifiers, it is easy to see that the mapping from relations to their names is a bijection from isomorphism classes of objects of Sub($B \times A$) to morphisms $B \to PA$.) We say $\mathcal{E}$ *has power objects* if there is an operation assigning to each object of $\mathcal{E}$ a power object for it. By the preceding discussion, a topos has power objects; conversely, it is clear that a cartesian category with power objects has a subobject classifier, since the latter is (up to isomorphism) the same thing as a power object for the terminal object 1. We shall see eventually that a cartesian category with power objects is also cartesian closed (in fact, locally cartesian closed) and hence a topos; but since we shall not arrive at this result until the middle of Section A2.3, we shall for the moment use the term *weak topos* for a cartesian category with power objects, in order to avoid confusion.

If $F: \mathcal{E} \to \mathcal{F}$ is a cartesian functor between (weak) toposes, we have for each object A of $\mathcal{E}$ a comparison morphism

$$\phi_A : F(PA) \longrightarrow P(F(A)),$$

which is the name of the relation $F(\in_A) \rightarrowtail F(PA \times A) \cong F(PA) \times FA$ (recall that a cartesian functor preserves monomorphisms as well as finite products). We say F is a *logical functor* if ϕ_A is an isomorphism for all A; equivalently, if $F(PA)$, equipped with the relation $F(\in_A)$, is a power object for $F(A)$ in $\mathcal{F}$, for all A. Once we have proved that a weak topos is a topos, it will follow that a logical functor is cartesian closed, as defined in Section A1.5; conversely, if F is cartesian closed and ϕ_1 is an isomorphism then F is logical, since ϕ_A may be factored as

$$F(\Omega_{\mathcal{E}}^A) \xrightarrow{\theta_{A,\Omega}} F(\Omega_{\mathcal{E}})^{F(A)} \xrightarrow{\phi_1^{F(A)}} \Omega_{\mathcal{F}}^{F(A)}$$

(where $\Omega_{\mathcal{E}}$ and $\Omega_{\mathcal{F}}$ denote the subobject classifiers of $\mathcal{E}$ and $\mathcal{F}$ respectively). Thus a cartesian functor is logical iff it 'preserves exponentials and the subobject classifier'.

We now turn to our list of examples.

Example 2.1.2 The category **Set** is of course a topos. So too is the full subcategory $\mathbf{Set}_f \subseteq \mathbf{Set}$ of finite sets (and the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$ is

logical); more generally, the category $\mathbf{Set}_\kappa$ of sets of cardinality less than κ is a topos provided κ is a limit power cardinal, i.e. a cardinal such that $\lambda < \kappa$ implies $2^\lambda < \kappa$.

Example 2.1.3 For any small category $\mathcal{C}$, the functor category $[\mathcal{C}, \mathbf{Set}]$ is a topos, as we saw in 1.5.5 and 1.6.6. Again, if $\mathcal{C}$ is a finite category, then $[\mathcal{C}, \mathbf{Set}_f]$ is a topos; the proofs of the results just cited still work with 'set' replaced by 'finite set' throughout. We shall see later (in Section B2.3) that if $\mathcal{C}$ is a finite category and $\mathcal{E}$ is any topos whatever, then $[\mathcal{C}, \mathcal{E}]$ is a topos.

Example 2.1.4 Let G be a group, i.e. a small groupoid with one object. A functor $G \to \mathbf{Set}$ is the same thing as a (left) G-*set*, i.e. a set on which G acts (associatively) by permutations. So, by 2.1.3, the category $[G, \mathbf{Set}]$ of G-sets (and G-equivariant maps) is a topos. It is worth describing the topos structure explicitly in this case: Ω is a two-element set with trivial G-action (cf. 1.4.12 and 1.6.5), and the exponential B^A is, by the construction of 1.5.5, the set of G-equivariant maps $G \times A \to B$ (where G acts on itself by left translations). But such a map $f: G \times A \to B$ is entirely determined by its restriction to $\{e\} \times A$, where e is the identity element of G, since $f(g, a) = g \cdot f(e, g^{-1} \cdot a)$; and this restriction may be any function $A \to B$. So we may re-interpret B^A as the set of all functions $A \to B$; the G-action is then by conjugation (i.e. $(g \cdot f)(a) = g \cdot (f(g^{-1} \cdot a))$. In particular, we note that the forgetful functor $[G, \mathbf{Set}] \to \mathbf{Set}$ is logical. It is also easy to verify that the above description of Ω and exponentials yields a topos structure on the category $[G, \mathbf{Set}_f]$ of finite G-sets, even when G is an infinite group; similarly, if $\mathcal{G}$ is a 'large group' (i.e. a one-object groupoid which is not small) then $[\mathcal{G}, \mathbf{Set}]$ is still a topos.

Example 2.1.5 Generalizing the last observation, let $\mathcal{C}$ be a category such that each slice category $\mathcal{C}/B$, $B \in \operatorname{ob} \mathcal{C}$, is equivalent to a small category, even though $\mathcal{C}$ itself may not be small. (Any groupoid $\mathcal{G}$ satisfies this condition, because any slice $\mathcal{G}/B$ is a trivial connected groupoid, and hence equivalent to $\mathbf{1}$; other examples of such categories include the ordered class $\mathbf{On}$ of ordinals, and the category $\mathbf{Set}_m$ of sets and injective functions.) Then it is not hard to see that (the duals of) the constructions of 1.5.5 and 1.6.6 'preserve smallness' (that is, they yield set-valued rather than class-valued functors), and hence that $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is a topos. (Similar examples can be constructed with 'finite' replacing 'small'.) We note that the toposes which arise in this way, unlike those we have considered hitherto, need not be locally small; for example, in $[\mathbf{On}^{\mathrm{op}}, \mathbf{Set}]$, there is a proper class of distinct subfunctors of (the constant functor with value) 1, and hence a proper class of morphisms $1 \to \Omega$.

Example 2.1.6 Generalizing in a different direction, let G be a topological group. We say that a G-set A is *continuous* if the action map $G \times A \to A$ is continuous for the discrete topology on A; equivalently, if for each $a \in A$ the

stabilizer subgroup (or isotropy subgroup)

$$G_a = \{g \in G \mid g \cdot a = a\}$$

is open in G. We shall show that the full subcategory $\mathbf{Cont}(G) \subseteq [G, \mathbf{Set}]$ of continuous G-sets is a topos. First we observe that the inclusion $\mathbf{Cont}(G) \to [G, \mathbf{Set}]$ has a right adjoint: given an arbitrary G-set A, the subset

$$A_c = \{a \in A \mid G_a \text{ is open in } G\}$$

is a union of G-orbits (because elements in the same G-orbit have conjugate stabilizers) and hence a sub-G-set of A; and the inclusion $A_c \to A$ is universal among morphisms from continuous G-sets to A (because, if $f: B \to A$ is a G-equivariant map and $b \in B$, the stabilizer $G_{f(b)}$ contains G_b, and hence is open if G_b is). Now $\mathbf{Cont}(G)$ is closed under finite limits in $[G, \mathbf{Set}]$ (and hence cartesian); it follows easily that we may construct exponentials in $\mathbf{Cont}(G)$ by constructing them in $[G, \mathbf{Set}]$ and then applying the functor $A \mapsto A_c$. Finally, the subobject classifier of $[G, \mathbf{Set}]$ lives in $\mathbf{Cont}(G)$; and since every sub-G-set of a continuous G-set is continuous, it is still a subobject classifier there.

There is some spurious generality in allowing G to be an arbitrary topological group in Example 2.1.6. Of course, if a connected topological group acts continuously on a discrete set, the action must be trivial; and so we may as well assume that G is totally disconnected, since for a general G we have $\mathbf{Cont}(G) \cong \mathbf{Cont}(G/G_0)$, where G_0 is the connected component of the identity element of G. In fact we can go slightly further than this, and restrict ourselves to groups which are *nearly discrete*, i.e. such that the intersection of all open subgroups is the trivial subgroup $\{e\}$ (again, if it isn't, we can factor it out without affecting $\mathbf{Cont}(G)$); the additive group of rationals is a totally disconnected group which is not nearly discrete. (This is a bad counterexample, since the uniform completion of the rationals is connected; there are complete counterexamples, but they are much harder to construct – see [1121].) A further possible restriction is to the class of those (nearly discrete) topological groups in which the open subgroups form a base of neighbourhoods of the identity; indeed, from one point of view it would be more sensible to describe this example as a construction on groups equipped with a normal filter of subgroups (i.e. one closed under conjugation), rather than with a topology. (Note that every normal filter of subgroups $\mathcal{F}$ generates a group topology, in which we take the cosets gH, $H \in \mathcal{F}$, as a base of neighbourhoods of an arbitrary $g \in G$. But a group topology is not uniquely determined by its filter of open subgroups, even if it is nearly discrete.)

However, we cannot cut down to the class of *pro-discrete* groups, i.e. groups which are expressible as inverse limits of discrete groups (equivalently, groups in which the intersection of all open normal subgroups is trivial). Two important examples of nearly discrete groups, which we shall encounter in Sections D3.4 and

C5.4, are the group of all permutations of $\mathbb{N}$ and the group of all order-preserving permutations of $\mathbb{Q}$ (topologized in each case by saying that the pointwise stabilizers of finite subsets of $\mathbb{N}$ (resp. $\mathbb{Q}$) form a base of neighbourhoods of the identity element), neither of which has any proper open normal subgroups.

The foregoing remarks form the beginnings of a theory of 'Morita equivalence' for topological groups, where we call two such groups *Morita-equivalent* if their toposes of continuous actions are equivalent. In B3.4.14(d) we shall generalize the result of 2.1.6 from topological groups to topological groupoids (and subsequently to localic groupoids); and in Section C5.3 we shall investigate the general notion of Morita equivalence for such groupoids.

Example 2.1.7 A seemingly minor, but significant, variation of Example 2.1.6: again, let G be a topological group, but this time consider *uniformly continuous* G-sets, i.e. those G-sets A for which the pointwise stabilizer of the whole of A (the subgroup of elements of G acting as the identity on A) is open in G. It is easy to see that if A and B are arbitrary G-sets, the pointwise stabilizer of $A \times B$ (resp. B^A) equals (resp. contains) the intersection of the pointwise stabilizers of A and B; hence the full subcategory $\mathbf{Unif}(G) \subseteq [G, \mathbf{Set}]$ of uniformly continuous G-sets is closed under finite limits and exponentials. It is also closed under arbitrary subobjects, and contains the subobject classifier of $[G, \mathbf{Set}]$; hence it is a topos, and the inclusion functor $\mathbf{Unif}(G) \to [G, \mathbf{Set}]$ is logical (as is the inclusion $\mathbf{Unif}(G) \to \mathbf{Cont}(G)$ – although $\mathbf{Cont}(G) \to [G, \mathbf{Set}]$ is not in general). In this case there is no loss of generality in restricting our attention to pro-discrete groups, since the pointwise stabilizer of a G-set is a normal subgroup of G; thus (for example) if G is either of the permutation groups mentioned in the last paragraph but one above, then $\mathbf{Unif}(G)$ is isomorphic to $\mathbf{Set}$. However, if we take G to be the profinite completion of the additive group $\mathbb{Z}$ of integers (i.e. the inverse limit of its finite quotients), then continuous G-sets may be identified with $\mathbb{Z}$-sets in which every orbit is finite, and such a set is uniformly continuous iff there is a finite bound for the sizes of its orbits; thus $\mathbf{Unif}(G)$ is not equivalent either to $\mathbf{Set}$ or to $\mathbf{Cont}(G)$. Note also that in this case $\mathbf{Unif}(G)$ does not have infinite coproducts, and hence cannot be coreflective in $[G, \mathbf{Set}]$.

Example 2.1.8 Let X be a topological space, and let $\mathcal{O}(X)$ denote the lattice of open sets of X, considered as a (small) preorder. A *presheaf* (of sets) on X is, by definition, a functor $\mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Set}$; thus the category $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ of presheaves on X is a topos. If F is a presheaf on X, we call the elements of $F(U)$ (for U an open set in X) *sections* of F over U, and we describe the map $F(U) \to F(V)$ induced by an inclusion $V \subseteq U$ in $\mathcal{O}(X)$ as *restriction* of sections from U to V. The example we have in mind is that, for any object $p: Y \to X$ of $\mathbf{Sp}/X$, we may define a presheaf $\Gamma(p)$ by taking $\Gamma(p)(U)$ to be the set of sections of p over U (i.e. continuous maps $s: U \to Y$ such that ps is the inclusion map $U \to X$); in fact, as is easily seen, Γ defines a functor $\mathbf{Sp}/X \to [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$. Presheaves in the image of this functor have the property that we can 'patch together' sections over the members of an open covering of U to obtain sections

over U: abstracting from this example, we define a presheaf F to be a *sheaf* if, given any open covering $(U_i \mid i \in I)$ of an open set U, and any family $(s_i \mid i \in I)$ of elements of $F(U_i)$ which are *compatible* in the sense that, for each pair (i, j), the restrictions of s_i and s_j are equal in $F(U_i \cap U_j)$, there exists a unique $s \in F(U)$ whose restriction to each U_i equals s_i. We write $\mathbf{Sh}(X)$ for the full subcategory of $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ whose objects are sheaves.

We claim that $\mathbf{Sh}(X)$ is a topos. The full proof of this is deferred until Section A4.4, where we shall in fact prove a much more general result; but we sketch the argument here for completeness. First, since limits in $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ are computed pointwise, $\mathbf{Sh}(X)$ is closed under (finite) limits (in fact it is reflective in $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$), and hence cartesian. Next, it is an exponential ideal in $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$: if G is a sheaf and F is any presheaf, then the exponential G^F, constructed as in 1.5.5, is a sheaf. Thus in particular $\mathbf{Sh}(X)$ is cartesian closed. Finally, we define a presheaf Ω on X by taking $\Omega(U)$ to be the set of open subsets of U, the restriction map $\Omega(U) \to \Omega(V)$ being $(-) \cap V$; this is easily verified to be a sheaf (patching is just forming unions in this case), and we have a map $\top\colon 1 \to \Omega$ defined by $\top_U = U \in \Omega(U)$. Now if $F' \rightarrowtail F$ is a monomorphism in $\mathbf{Sh}(X)$ (for convenience we assume that the values $F'(U)$ of F' are actual subsets of $F(U)$) and $s \in F(U)$, there is a unique largest $V \subseteq U$ such that the restriction of s to V belongs to $F'(V)$ (namely the union of all such V), because F' is a sheaf; so defining $h_U(s)$ to be this V produces a morphism $h\colon F \to \Omega$, which may be verified to be the classifying map of $F' \rightarrowtail F$.

In Section C1.3 we shall show that the functor $\Gamma\colon \mathbf{Sp}/X \to [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ defined in the preceding example restricts to (one half of) an equivalence $\mathbf{LH}/X \simeq \mathbf{Sh}(X)$, so that $\mathbf{LH}/X$ is also a topos. However, we shall provide an independent proof of the latter fact in 4.2.4(*e*) below.

It is worth noting from Examples 2.1.6 and 2.1.8 that the properties of preserving Ω and preserving exponentials are independent, even for cartesian functors which are full and faithful: the inclusion $\mathbf{Cont}(G) \to [G, \mathbf{Set}]$ preserves Ω but not exponentials, and the inclusion $\mathbf{Sh}(X) \to [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ preserves exponentials but not Ω.

Generalizing from the last example, we may define a notion of sheaf on any small category $\mathcal{C}$, provided it is equipped with a suitable notion of 'covering', as follows:

Definition 2.1.9 Let $\mathcal{C}$ be a small category. By a *coverage* on $\mathcal{C}$, we mean a function assigning to each object A of $\mathcal{C}$ a collection $T(A)$ of families $(f_i\colon A_i \to A \mid i \in I)$ of morphisms with common codomain A (called *T-covering families*), such that

(*C*) If $(f_i\colon A_i \to A \mid i \in I)$ is a T-covering family and $g\colon B \to A$ is any morphism with codomain A, there exists a T-covering family $(h_j\colon B_j \to B \mid j \in J)$ such that each gh_j factors through some f_i.

A *site* is a small category equipped with a coverage.

The old name for coverage is 'Grothendieck (pre-)topology'; but since this concept is a very considerable abstraction from the notion of a topological space (and, in particular, has nothing to do with the notion of a topology on a set), the latter name seems unhelpful. Also, the definition of a Grothendieck topology is normally taken to include additional 'stability properties' of covering families, besides (C) above, which are satisfied in many examples but are really irrelevant to the definition of a sheaf (we shall meet these conditions in Section C2.1, when we introduce the notion of *Grothendieck coverage*); and that of a Grothendieck pretopology normally includes the additional assumption that the underlying category $\mathcal{C}$ has pullbacks, and replaces (C) by the stronger condition

(C') If $(f_i : A_i \to A \mid i \in I)$ is a T-covering family and $g: B \to A$ is any morphism with codomain A, then the family $(g^*(f_i) \mid i \in I)$ is T-covering.

Again, (C') is satisfied in most of the examples we shall meet, but it is not necessary for the purpose of defining sheaves, and the extra generality allowed by (C) will be useful in some cases.

We say a functor $F: \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ *satisfies the sheaf axiom* for a family of morphisms $(f_i: A_i \to A \mid i \in I)$ if, whenever we are given a family of elements $s_i \in F(A_i)$ which are *compatible* in the sense that, whenever $g: B \to A_i$ and $h: B \to A_j$ satisfy $f_i g = f_j h$ (here i and j need not be distinct), we have $F(g)(s_i) = F(h)(s_j)$, then there exists a unique $s \in F(A)$ such that $F(f_i)(s) = s_i$ for each $i \in I$. (Once again, this definition may be simplified if $\mathcal{C}$ has pullbacks; it then suffices to check the compatibility of s_i and s_j on the pullback of f_i against f_j, rather than on arbitrary pairs (g, h) as above. In this case, the condition may conveniently be expressed diagrammatically: F satisfies the sheaf axiom for $(A_i \to A \mid i \in I)$ iff

$$F(A) \longrightarrow \prod_{i \in I} F(A_i) \rightrightarrows \prod_{i,j} F(A_i \times_A A_j)$$

is an equalizer diagram.) We say F is a *T-sheaf* if it satisfies the sheaf axiom for every T-covering family; and we write $\mathbf{Sh}(\mathcal{C}, T)$ (or simply $\mathbf{Sh}(\mathcal{C})$, if T is obvious from the context) for the full subcategory of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ whose objects are T-sheaves.

Lemma 2.1.10 *If $(\mathcal{C}, T)$ is a site, then $\mathbf{Sh}(\mathcal{C}, T)$ is a topos.*

Proof Again, we shall only sketch the details here; for a more detailed proof, we refer to 4.4.5 below, or to C2.2.6 (where we shall prove a more general result, under a weaker hypothesis on $\mathcal{C}$ than smallness). The fact that $\mathbf{Sh}(\mathcal{C}, T)$ is closed under finite limits (indeed, under arbitrary limits) in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is immediate from the form of the definition. It is also an exponential ideal, and hence cartesian closed. We recall from 1.6.6 that the subobject classifier of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ was

defined by

$$\Omega(A) = \{\text{sieves on } A \text{ in } \mathcal{C}\} .$$

We define a sieve R on A to be *T-closed* if, given $f: B \to A$ in $\mathcal{C}$, the existence of a T-covering family $(g_j: B_j \to B \mid j \in J)$ such that each fg_j belongs to R forces $f \in R$; and we write $\Omega_T(A) \subseteq \Omega(A)$ for the set of T-closed sieves. It then follows easily from condition (C) that Ω_T is a subfunctor of Ω; moreover, every sieve has a T-closure (the intersection of the T-closed sieves which contain it), and using this we may define a notion of patching for closed sieves and hence show that Ω_T is a sheaf. Finally, if F is a sheaf, it follows easily from the proof of 1.6.6 that the classifying map $F \to \Omega$ of a subfunctor $F' \rightarrowtail F$ factors through $\Omega_T \rightarrowtail \Omega$ iff F' is a sheaf; hence Ω_T is a subobject classifier for $\mathbf{Sh}(\mathcal{C}, T)$. □

It is clear that if $\mathcal{C} = \mathcal{O}(X)$ for a space X and we take $T(U)$ to consist of all jointly-surjective families of open inclusions $(U_i \subseteq U \mid i \in I)$, then the notion of T-sheaf coincides with that of a sheaf on X; so 2.1.10 includes 2.1.8 as a special case. But there are many other examples of toposes which can be constructed as $\mathbf{Sh}(\mathcal{C}, T)$ for suitable sites $(\mathcal{C}, T)$, and which will play important rôles in various parts of this book. We list some of them here: however, we should remark that for a given topos $\mathcal{E}$ there may be many different *sites of definition* for $\mathcal{E}$ (that is, sites $(\mathcal{C}, T)$ such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, T)$), and the sites which we describe below are not necessarily the most convenient ones for defining these particular toposes. In C2.2.3 we shall prove a 'comparison lemma' which enables us to determine when two sites give rise to equivalent toposes of sheaves.

Examples 2.1.11 (*a*) Let $\mathcal{C}$ be a small regular category. The *regular coverage* on $\mathcal{C}$ has, as covering families on A, all singletons $\{f\}$ where f is a cover (in the sense of Section A1.3) with codomain A. Condition (C') is satisfied by the very definition of a regular category. We note that Proposition 1.3.4 has an interpretation in terms of this coverage: it says that every representable functor $\mathcal{C}(-, A): \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is a sheaf for it (since such functors take coequalizers in $\mathcal{C}$ to equalizers in $\mathbf{Set}$). (A site is called *standard* if its underlying category $\mathcal{C}$ is cartesian and all the representable functors $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ are sheaves; we shall see in C2.2.8 that every topos generated by a site can be generated by a standard one, and in fact most of the sites one encounters in practice are standard.) Thus the Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ factors through $\mathbf{Sh}(\mathcal{C})$; the resulting functor $Y: \mathcal{C} \to \mathbf{Sh}(\mathcal{C})$ is still full, faithful and cartesian, but it also preserves covers (in general, if $f: A \to B$ is a morphism of $\mathcal{C}$, the image in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ of $Y(f): \mathcal{C}(-, A) \to \mathcal{C}(-, B)$ is (the subfunctor of $\mathcal{C}(-, B)$ corresponding to) the sieve on B of all morphisms factoring through f, and if f is a cover it is easy to see that this cannot be contained in any proper subsheaf of $\mathcal{C}(-, B)$). Thus we conclude that *any small regular category admits a full and faithful regular functor into a topos.* It is worth noting, too, that if $\mathcal{C}$ happens to be cartesian closed as well as regular, then this structure is also preserved by the embedding

$\mathcal{C} \to \mathbf{Sh}(\mathcal{C}, T)$, by 1.5.6(i) and the fact that $\mathbf{Sh}(\mathcal{C}, T)$ is an exponential ideal in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

(*b*) Let $\mathcal{C}$ be a small coherent category. The *coherent coverage* on $\mathcal{C}$ similarly has as covering families all finite families $(f_i \colon A_i \to A \mid 1 \leq i \leq n)$ which are covering in the sense defined in Section A1.3 – equivalently, such that the union of the images of the f_i is the whole of A. (In particular, the initial object 0 of $\mathcal{C}$ is covered by the empty family.) Once again, this site is standard; the extra information which we need (in addition to 1.3.4) to show that the representable functors are sheaves is contained in 1.4.1 and 1.4.3. (The toposes which arise from coherent sites are called *coherent toposes* (unsurprisingly, but perhaps unfortunately – as we shall see, *every* topos is a coherent category); they will be investigated in Section D3.3.) And once again, it may be verified that the Yoneda embedding, as a functor $\mathcal{C} \to \mathbf{Sh}(\mathcal{C})$, preserves covers and finite unions, and hence *any small coherent category admits a full and faithful coherent functor into a topos.*

(*c*) The next example is a synthesis of Examples 2.1.4 and 2.1.8. Let G be a (discrete) group, and X a space on which G acts by homeomorphisms. The topos $\mathbf{Sh}_G(X)$ of *G-equivariant sheaves* on X is defined to be $\mathbf{Sh}(\mathcal{O}_G(X), T)$, where the category $\mathcal{O}_G(X)$ may conveniently be defined by a protocategory, as in Section A1.1: its objects are the open subsets of X, its protomorphisms are the elements of G, and its source–target predicate says that $g \colon V \to U$ iff the homeomorphism $\overline{g} \colon X \to X$ induced by g maps V into U. Composition is induced by composition in G (that is, $\mathcal{O}_G(X)$ is a category structured over G, as defined in 1.1.5). The coverage T is defined by saying that a family $(g_i \colon V_i \to U \mid i \in I)$ is covering iff the union of the images of the $\overline{g_i} \colon V_i \to U$ is the whole of U. It is not hard to see that $\mathcal{O}_G(X)$ has pullbacks, and that T satisfies condition (C'). If Y is another space on which G acts, and $p \colon Y \to X$ is a G-equivariant continuous map, then the assignment $U \mapsto \Gamma(p)(U)$, defined in 2.1.8, extends to a functor $\mathcal{O}_G(X)^{\mathrm{op}} \to \mathbf{Set}$: if $s \colon U \to Y$ is a section of p over U, and $g \colon V \to U$ in $\mathcal{O}_G(X)$, then the composite

$$V \xrightarrow{\ \overline{g}\ } U \xrightarrow{\ s\ } Y \xrightarrow{\ \overline{g^{-1}}\ } Y$$

is a section of p over V, which we define to be $\Gamma(p)(g)(s)$. It is also easy to verify that $\Gamma(p)$ is a T-sheaf, and that Γ defines a functor $[G, \mathbf{Sp}]/X \to \mathbf{Sh}_G(X)$, where $[G, \mathbf{Sp}]$ is the category of spaces equipped with a G-action and G-equivariant continuous maps. Once again, this functor becomes part of an equivalence if we restrict it to the full subcategory $[G, \mathbf{LH}]/X \subseteq [G, \mathbf{Sp}]/X$ whose objects are G-spaces equipped with a G-equivariant local homeomorphism into X.

(*d*) Let $\mathcal{C}$ be a small full subcategory of $\mathbf{Sp}$ which is closed under passage to open subspaces (at least in the sense that, if X is in $\mathcal{C}$ and U is an open subspace of X, then some space homeomorphic to U is in $\mathcal{C}$). For example, we might take (a suitable skeleton of) the category of all topological spaces of cardinality less than κ, for a suitable cardinal κ. On $\mathcal{C}$ we have a coverage defined by saying that

the covering families are open coverings in the usual sense, i.e. jointly-surjective families of open inclusions. (Condition (C') is satisfied because the pullback of an open inclusion along an arbitrary continuous map is an open inclusion.) Once again, this site is standard (at least if $\mathcal{C}$ is closed under finite limits in **Sp**): the assertion that the representable functors are sheaves is essentially the statement that, if X and Y are spaces and $f: X \to Y$ is a function which is continuous when restricted to each set in some open covering of X, then f is continuous. This example can also be given in a G-equivariant version (where G is a fixed group), similar to (c) above.

(*e*) Similar examples to (*d*) may be constructed from appropriate full subcategories of the category **Mf** of (paracompact) smooth manifolds and smooth maps, again taking open coverings to define the coverage. A similar assertion to that at the end of the last paragraph holds with 'smooth' replacing 'continuous'; so once again the representable functors on such a site are sheaves. However, the category **Mf** does not have all pullbacks (for example, if $f: \mathbb{R} \to \mathbb{R}$ is a smooth function such that $f(x) \neq 0$ iff $x > 0$, then the functions $x \mapsto (x, 0)$ and $x \mapsto (x, f(x))$ from $\mathbb{R}$ to $\mathbb{R}^2$ have no pullback); so this is one example where we obtain the benefit of not insisting, in our definition of a site, that the underlying category should have pullbacks. (On the other hand, the pullback of an open inclusion along an arbitrary morphism does exist in **Mf**; and the strong condition (C') is still satisfied in this case.)

(*f*) There are also 'algebraic' versions of the last two examples, in which the category of spaces or manifolds is replaced by the category of affine schemes (the opposite of the category of commutative rings). For definiteness, let k be a field, and let $\mathcal{C}^{\mathrm{op}}$ be the category of finitely-presented k-algebras (= quotients of polynomial rings in finitely many variables over k). The *Zariski coverage* Z on $\mathcal{C}$ consists of jointly-surjective families of Zariski-open inclusions; in algebraic terms, these are families of homomorphisms of the form

$$(A \longrightarrow A[s^{-1}] \mid s \in S)$$

where $A \in \mathrm{ob}\,\mathcal{C}$ and S is a set of elements of A which is not contained in any proper ideal of A. For any homomorphism $f: A \to B$ of k-algebras, we have a pushout square

$$\begin{tikzcd} A \arrow[r] \arrow[d, "f"] & A[s^{-1}] \arrow[d] \\ B \arrow[r] & B[(f(s))^{-1}] \end{tikzcd}$$

in $\mathcal{C}^{\mathrm{op}}$, from which it is easy to verify that the Zariski coverage satisfies condition (C'). The topos $\mathbf{Sh}(\mathcal{C}, Z)$ is called the *Zariski topos* over k.

(*g*) Of course, the open-set lattice $\mathcal{O}(X)$ of a space X, with the usual notion of open covering, provides an example of a site whose underlying category is

a poset. But there are other such examples: we describe a particular one here, which we shall meet again in C3.3.11(c) and D4.1.9. Let A and B be two fixed infinite sets, and let P be the set of partial functions $A \rightharpoondown B$ with finite domain (that is, functions from finite subsets of A to B), ordered by the relation 'is an extension of' (i.e., $f \leq g$ if dom $g \subseteq$ dom f and $g(a) = f(a)$ whenever $g(a)$ is defined). We define a coverage T by saying that $T(f)$ consists of all sets of the form

$$\{g \in P \mid g \leq f \text{ and } b \in \text{im } g\}$$

where b ranges over the elements of B. (Note that here, as we usually do when dealing with coverages on posets, we are thinking of covers of f as sets of elements of P lying below f, rather than as families of morphisms with codomain f.) It is again an easy exercise to verify that the representable functors $P^{\text{op}} \to \mathbf{Set}$ are all sheaves; but (P, T) is not a standard site, because P does not have pullbacks (and here the use of the weak form (C) of the stability condition is essential: it is also easy to verify, since if $(g_i \mid i \in I)$ is the cover of f associated with a particular element b of B, and $h \leq f$, then the cover of h associated with the same element b satisfies the condition).

An important property of this particular site is that *the disjoint union (i.e. the coproduct in* $[P^{\text{op}}, \mathbf{Set}]$*) of any family of sheaves is again a sheaf*; in general the category $\mathbf{Sh}(\mathcal{C}, T)$ is not closed under coproducts (or any other colimits) in $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$. In particular, since the constant functor with value 1 is a sheaf, we see that *all constant functors* $P^{\text{op}} \to \mathbf{Set}$ *are sheaves.* To verify this property, suppose we have a family of sheaves $(F_i \mid i \in I)$, a particular cover $(g_j \mid j \in J) \in T(f)$, and a compatible family $(x_j \mid j \in J)$ of elements of the disjoint unions $\coprod_{i \in I} F_i(g_j)$, $j \in J$. We claim that the x_j must in fact all belong to $F_i(g_j)$ for the same index i, so that they can be uniquely 'patched together' to an element $x \in F_i(f) \subseteq \coprod_{i \in I} F_i(f)$. For if g_j and $g_{j'}$ satisfy $g_j \leq g_{j'} \leq f$, then it is clear that x_j and $x_{j'}$ must belong to the same summand of the coproduct; and for any two elements g_1 and g_2 we can find a zigzag

$$g_1 \leq g_3 \geq g_5 \leq g_4 \geq g_2$$

of elements of the covering family linking them. Specifically, suppose the cover is that associated with an element b of B; we assume b is not in the image of f (if it is, then f itself belongs to the cover and we have nothing to prove). Then we take g_3 to be an extension of f defined at just one more element a_1 of A than f, chosen so that $g_1(a_1) = b$; and g_4 is similarly defined from g_2. If g_3 and g_4 are not identical (because the new elements a_1, a_2 of their domains are not equal) then we take g_5 to be their unique common extension with domain dom $f \cup \{a_1, a_2\}$. (We shall return to this 'zigzag condition' on covers in C3.3.10.)

(h) Given a category $\mathcal{C}$, it is easy to see that the collection of all singleton families (f), f a morphism of $\mathcal{C}$, satisfies the coverage axiom (C) iff, given any two morphisms $f\colon A \to C$, $g\colon B \to C$ of $\mathcal{C}$ with common codomain, we can find

morphisms h, k forming a commutative square:

$$\begin{tikzcd} D \arrow[r, "h"] \arrow[d, "k"'] & A \arrow[d, "f"] \\ B \arrow[r, "g"] & C \end{tikzcd} .$$

For monoids, this condition is commonly known as the *right Ore condition*, and it seems sensible to use this terminology for categories as well. It is clearly implied by the condition that C has pullbacks (in which case the coverage satisfies the strong condition (C')), but is considerably weaker. There are several interesting examples of sites where all singletons cover; one which will reappear at several points in this book is that obtained by taking C to be the *dual* of the category $\mathbf{Set}_{fm}$ of finite sets and monomorphisms. The corresponding topos $\mathbf{Sh}(\mathbf{Set}_{fm}^{\mathrm{op}}, T)$ is commonly known as the *Schanuel topos*, in honour of S. H. Schanuel who was the first to investigate its properties; it was also extensively exploited by A. Joyal [552].

For the present, we note an interesting characterization of the objects of the Schanuel topos: they are exactly the functors $\mathbf{Set}_{fm} \to \mathbf{Set}$ which preserve pullbacks. (Note that $\mathbf{Set}_{fm}$ has pullbacks – indeed, the inclusion $\mathbf{Set}_{fm} \to \mathbf{Set}_f$ creates them – though it does not have a terminal object.) To see this, note first that a functor $F\colon \mathbf{Set}_{fm} \to \mathbf{Set}$ satisfies the sheaf axiom for the singleton family $(f\colon [m] \rightarrowtail [n])$ (where we use the notation $[n]$ for an n-element finite set) iff $F(f)$ is injective, and its image consists precisely of those $x \in F([n])$ satisfying $F(g)(x) = F(h)(x)$ whenever $gf = hf$. But the first condition is implied by the assertion that F preserves the pullback of f against itself; for the second, we can always find a pair (g, h) such that

$$\begin{tikzcd} {[m]} \arrow[r, tail, "f"] \arrow[d, tail, "f"'] & {[n]} \arrow[d, tail, "g"] \\ {[n]} \arrow[r, tail, "h"] & {[p]} \end{tikzcd}$$

is a pullback, so the condition that F preserves pullbacks implies that if $F(g)(x) = F(h)(x)$ then there exists $y \in F([m])$ with $F(f)(y) = x$. Conversely,

if F is assumed to be a sheaf, and we are given an arbitrary pullback square

$$\begin{tikzcd}
{[m]} \arrow[r, tail, "f"] \arrow[d, tail, "g"'] & {[n]} \arrow[d, tail, "h"] \\
{[p]} \arrow[r, tail, "k"] & {[q]}
\end{tikzcd} ,$$

then given elements $x \in F([n]), y \in F([p])$ with $F(h)(x) = F(k)(y)$, we claim that for any pair of morphisms (a, b) with $af = bf$ we can extend the above diagram to one of the form

$$\begin{tikzcd}
{[m]} \arrow[r, tail, "f"] \arrow[d, tail, "g"'] & {[n]} \arrow[r, tail, shift left, "a"] \arrow[r, tail, shift right, "b"'] \arrow[d, tail, "h"] & {[r]} \arrow[d, tail, "e"] \\
{[p]} \arrow[r, tail, "k"] & {[q]} \arrow[r, tail, shift left, "c"] \arrow[r, tail, shift right, "d"'] & {[s]}
\end{tikzcd}$$

with $ea = ch$, $eb = dh$ and $ck = dk$ (just take $s = r + q - n$), and so we necessarily have $F(e)F(a)(x) = F(c)F(k)(y) = F(d)F(k)(y) = F(e)F(b)(x)$, whence $F(a)(x) = F(b)(x)$ since $F(e)$ is injective. Thus there exists a unique $z \in F([m])$ with $F(f)(z) = x$; and we also have $F(g)(z) = y$ since $F(k)$ is injective. So F preserves the above pullback.

(*i*) It is easily verified that if a covering family contains a split epimorphism, then every functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ satisfies the sheaf axiom for it. Hence a group does not admit any nontrivial coverage: if T is a coverage on a group G, then $\mathbf{Sh}(G, T)$ is either the whole of $[G^{\mathrm{op}}, \mathbf{Set}]$ (if every T-covering family is nonempty) or equivalent to $\mathbf{1}$ (if the empty family covers the unique object of G). However, monoids can admit nontrivial coverages. For example, let M be the free monoid on two generators x and y, and let T be the coverage whose only covering family (on the unique object $*$ of M) is $(x\colon * \to *, y\colon * \to *)$. The fact that T satisfies condition (C) is immediate from the fact that every morphism of M, other than the identity, factors through either x or y (that is, every nontrivial word in x and y must begin with either x or y). It is easy to see that x and y are monic in M, and that there are no pairs (z, t) with $xz = yt$; hence a T-sheaf is just a right M-set A (that is, a functor $M^{\mathrm{op}} \to \mathbf{Set}$) for which the map $a \mapsto (a \cdot x, a \cdot y)$ is a bijection $A \to A \times A$. The category $\mathbf{Sh}(M, T)$ is sometimes called the *Jónsson–Tarski topos*, in honour of B. Jónsson and A. Tarski who were the first to investigate it (as a variety of universal algebras, not as a topos) in [551]; see also [525].

(j) Another monoid-based example: let $\mathbb{N}^+$ be the one-point compactification of the discrete space $\mathbb{N}$ of natural numbers, and let M be the monoid of continuous endomorphisms of $\mathbb{N}^+$ (i.e. the full subcategory of **Sp** whose only object is $\mathbb{N}^+$). We say that a continuous map $f\colon \mathbb{N}^+ \to \mathbb{N}^+$ is *non-degenerate* if its image is infinite (which implies that f maps the non-isolated point ∞ of $\mathbb{N}^+$ to itself). We now define a family of morphisms $(f_i\colon \mathbb{N}^+ \to \mathbb{N}^+ \mid i \in I)$ to be covering if

(i) it is jointly surjective, and

(ii) for any infinite subset S of $\mathbb{N}$, there exists $i \in I$ such that f_i is non-degenerate and has image contained in $S \cup \{\infty\}$.

The verification that this defines a coverage on M is straightforward. Now, for any space X, the set $\mathbf{Sp}(\mathbb{N}^+, X)$ of continuous maps $\mathbb{N}^+ \to X$ (that is, convergent sequences in X, together with their limits) has an obvious right M-set structure by composition of continuous maps; and it is a sheaf for the coverage just described. (The key point in verifying this is the well-known fact that if (x_n) is a sequence of points of X and x_∞ a point such that every subsequence of (x_n) contains a subsequence converging to x_∞, then (x_n) converges to x_∞.) In fact the assignment $X \mapsto \mathbf{Sp}(\mathbb{N}^+, X)$ defines a functor $\mathbf{Sp} \to \mathbf{Sh}(M)$; this functor is faithful, and it is full when restricted to the subcategory of sequential spaces (i.e. spaces where convergent sequences determine the topology, in the sense that any subset closed under limits of sequences is actually closed; note that this subcategory contains all metrizable spaces). More detailed information about this example will be found in [507].

(k) Our final example, also monoid-based, is a 'recursive' as opposed to 'continuous' version of (j), due to P. S. Mulry [864]. Let M be the monoid of all (total) recursive functions $\mathbb{N} \to \mathbb{N}$, and define T to consist of all *finite* families $(f_i\colon \mathbb{N} \to \mathbb{N} \mid 1 \leq i \leq n)$ such that the union of the images of the f_i is the whole of $\mathbb{N}$. To verify condition (C) in this case, let $(f_i \mid 1 \leq i \leq n)$ be such a family, and $g\colon \mathbb{N} \to \mathbb{N}$ an arbitrary recursive function. For each i, the function

$$n \mapsto \mu m \,.\, f_i(m) = g(n)$$

(where $\mu m \,.\, P$, as usual, denotes 'the least m such that P') is partial recursive; and so if its domain $g^{-1}(\mathrm{im}\, f_i)$ is nonempty we can find a total recursive function h_i which enumerates it. Now the family $(h_i \mid g^{-1}(\mathrm{im}\, f_i) \neq \emptyset)$ is T-covering, and for each such i the function

$$k_i(n) = \mu m \,.\, f_i(m) = gh_i(n)$$

is total recursive and satisfies $f_i k_i = gh_i$. It is easy to verify that the unique representable functor $M^{\mathrm{op}} \to \mathbf{Set}$ is a T-sheaf (though it would not be so if we removed the requirement that members of T be finite); more generally, the category **En** of *enumerated sets* in the sense of Ershov [336] may be identified with a full subcategory of $\mathbf{Sh}(M, T)$. An enumerated set is a pair (A, e), where A is

a set and $e\colon \mathbb{N} \twoheadrightarrow A$ is a surjection; a morphism $f\colon (A,e) \to (B,d)$ of enumerated sets is a function $f\colon A \to B$ which is 'tracked' by some recursive function, i.e. for which there exists a recursive $g\colon \mathbb{N} \to \mathbb{N}$ such that

$$\begin{tikzcd} \mathbb{N} \arrow[r, "g"] \arrow[d, two heads, "e"] & \mathbb{N} \arrow[d, two heads, "d"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

commutes. We define a functor $F\colon \mathbf{En} \to \mathbf{Sh}(M,T)$ as follows: $F(A,e)$ is the set of all functions $\mathbb{N} \to A$ which can be written in the form eh for some total recursive h, with M acting by composition, and if $f\colon (A,e) \to (B,d)$ in $\mathbf{En}$, then $F(f)(eh) = feh$ ($= dgh$, where g is some recursive function tracking f). To show that every M-equivariant map $F(A,e) \to F(B,d)$ is induced by composition with a (unique) function $f\colon A \to B$, we use the fact that the constant maps $\mathbb{N} \to A$ may be identified as the M-invariant elements of $F(A,e)$; and this f must be a morphism of $\mathbf{En}$, because the composite fe belongs to $F(B,d)$. For more information about the properties of this topos, see [864].

We conclude this section with a couple of methods of constructing new toposes from old, which are often useful.

Example 2.1.12 [**the glueing construction**] Let $F\colon \mathcal{E} \to \mathcal{F}$ be a cartesian functor between toposes. We define the *glued topos* of F, $\mathbf{Gl}(F)$, by taking its objects to be triples (A,B,α) where A and B are objects of $\mathcal{E}$ and $\mathcal{F}$ respectively and $\alpha\colon B \to F(A)$ in $\mathcal{F}$, and morphisms $(A,B,\alpha) \to (C,D,\gamma)$ to be pairs $(f\colon A \to C, g\colon B \to D)$ such that

$$\begin{tikzcd} B \arrow[r, "\alpha"] \arrow[d, "g"] & F(A) \arrow[d, "F(f)"] \\ D \arrow[r, "\gamma"] & F(C) \end{tikzcd}$$

commutes. (In other words, $\mathbf{Gl}(F)$ is just the *comma category* $(\mathcal{F} \downarrow F)$, as usually defined; cf. B1.1.4(c).) It is immediate that $\mathbf{Gl}(F)$ is a category; also that it is cartesian, with finite limits constructed pointwise. We claim that *for any cartesian functor F between toposes,* $\mathbf{Gl}(F)$ *is a topos.* To construct the exponential

$(C, D, \gamma)^{(A,B,\alpha)}$, we form the pullback

$$\begin{tikzcd} E \arrow[rr, "k"] \arrow[d, "h"'] & & D^B \arrow[d, "\gamma^B"] \\ F(C^A) \arrow[r, "\theta"] & F(C)^{F(A)} \arrow[r, "F(C)^\alpha"] & F(C)^B \end{tikzcd}$$

(where θ is the exponential comparison morphism defined in Section A1.5), and we take $(C, D, \gamma)^{(A,B,\alpha)} = (C^A, E, h)$; then the evaluation morphism $(C^A \times A, E \times B, h \times \alpha) \to (C, D, \gamma)$ is $(\mathrm{ev}, \mathrm{ev}(k \times 1_B))$. Given a morphism $(f, g)\colon (X, Y, \xi) \times (A, B, \alpha) \to (C, D, \gamma)$, its exponential transpose $(X, Y, \xi) \to (C^A, E, h)$ has $\bar{f}$ as its first component and the morphism induced by $\bar{g}\colon Y \to D^B$ and $F(\bar{f}) \circ \xi\colon Y \to F(C^A)$ as its second. For the subobject classifier in $\mathbf{Gl}(F)$, we take $(\Omega_{\mathcal{E}}, Q, q)$, where Q is defined by the pullback square

$$\begin{tikzcd} Q \arrow[r] \arrow[d, tail, "(p,q)"'] & (\Omega_1)_{\mathcal{F}} \arrow[d, tail] \\ \Omega_{\mathcal{F}} \times F(\Omega_{\mathcal{E}}) \arrow[r, "1 \times \phi"] & \Omega_{\mathcal{F}} \times \Omega_{\mathcal{F}} \end{tikzcd} .$$

Given a subobject $(m, n)\colon (A', B', \alpha') \rightarrowtail (A, B, \alpha)$, the classifying map $(f, g)\colon (A, B, \alpha) \to (\Omega_{\mathcal{E}}, Q, q)$ of (m, n) is obtained by taking f to be the classifying map of m, and g to be the factorization through $Q \rightarrowtail \Omega_{\mathcal{F}} \times F(\Omega_{\mathcal{E}})$ of the composite

$$B \xrightarrow{(1,\alpha)} B \times F(A) \xrightarrow{h \times F(f)} \Omega_{\mathcal{F}} \times F(\Omega_{\mathcal{E}})$$

where h is the classifying map of n. The rest is straightforward verification.

We note that there are projection functors $P_1\colon \mathbf{Gl}(F) \to \mathcal{E}$ and $P_2\colon \mathbf{Gl}(F) \to \mathcal{F}$; the former is logical, but the latter is merely cartesian (unless F is (isomorphic to) the constant functor with value 1, in which case $\mathbf{Gl}(F)$ reduces to $\mathcal{E} \times \mathcal{F}$). Both these functors have full and faithful right adjoints: the right adjoint of P_1 sends A to $(A, F(A), 1_{F(A)})$, and the right adjoint of P_2 sends B to $(1, B, B \to 1)$.

Many instances of the glueing construction will be of use to us later on; one in particular is worth mentioning now. If $\mathcal{E}$ is a locally small topos, then the representable functor $\mathcal{E}(1, -)\colon \mathcal{E} \to \mathbf{Set}$ is cartesian: the topos obtained by glueing along it is called the *scone* (or Sierpiński cone) on $\mathcal{E}$, and denoted $\mathbf{scn}(\mathcal{E})$. In particular if $\mathcal{E}$ itself is $\mathbf{Set}$, then $\mathcal{E}(1, -)$ is (isomorphic to) the identity functor, and $\mathbf{scn}(\mathbf{Set})$ is just another name for the functor category $[\mathbf{2}, \mathbf{Set}]$ (equivalently,

for the topos of sheaves on the *Sierpiński space*, i.e. the two-point space with just one closed point). But other simple examples of toposes have more interesting scones: for example, $\mathbf{scn}(\mathbf{Set}_f)$ will be a fruitful source of counterexamples later on (see, for instance, 4.2.4(d) and 4.5.24 below). The scone construction will also be of pivotal importance in Section F3.2.

Example 2.1.13 [the filterquotient construction] Let $\mathcal{E}$ be a topos, and let Φ be a filter of subobjects of 1 in $\mathcal{E}$, that is, a collection of subterminal objects with the property that $U \in \Phi$ iff U contains the intersection of some finite set of members of Φ. (Incidentally, we shall often use letters such as U and V to denote subterminal objects in a topos; this is motivated by the example of $\mathbf{Sh}(X)$, where the subterminal objects correspond to open subsets of X.) By a *Φ-map* $A \rightharpoonup B$ in $\mathcal{E}$, we shall mean a morphism $A \times U \to B$ for some $U \in \Phi$; we say that two Φ-maps $f\colon A \times U \to B$ and $g\colon A \times V \to B$ are *equivalent* if there exists $W \leq U \cap V$ in Φ such that

$$\begin{tikzcd} A \times W \arrow[r, tail] \arrow[d, tail] & A \times U \arrow[d, "f"] \\ A \times V \arrow[r, "g"] & B \end{tikzcd}$$

commutes. Since Φ is closed under finite intersections, it is easy to see that this is indeed an equivalence relation. Φ-maps can be composed: given $f\colon A \times U \to B$ and $g\colon B \times V \to C$, we have a pullback square

$$\begin{tikzcd} A \times U \times V \arrow[r, "f \times 1"] \arrow[d, tail] & B \times V \arrow[d, tail] \\ A \times U \arrow[r, "f"] & B \end{tikzcd}$$

where $U \times V = U \cap V \in \Phi$, and so we may define the composite of g and f to be $g(f \times 1_V)$. This composition respects equivalence and is associative up to equivalence, so we obtain a category $\mathcal{E}_\Phi$, the *filterquotient* of $\mathcal{E}$ modulo Φ, whose objects are those of $\mathcal{E}$ but whose morphisms are equivalence classes of Φ-maps. And we have a functor $P_\Phi\colon \mathcal{E} \to \mathcal{E}_\Phi$, which is the identity on objects and sends $f\colon A \to B$ to the equivalence class of f regarded as a morphism $A \times 1 \to B$. We claim that *$\mathcal{E}_\Phi$ is a topos, and the functor P_Φ is logical.* In proving this, the key observation is that, since Φ is closed under finite intersections, any finite collection of morphisms $A_i \to B_i$ in $\mathcal{E}_\Phi$ may be simultaneously represented by morphisms $A_i \times U \to B_i$ in $\mathcal{E}$, for a single $U \in \Phi$. Using this, it is easy to see that $A \times B$ is still the product of A and B in $\mathcal{E}_\Phi$; and the equalizer of a parallel pair

of morphisms $A \rightrightarrows B$ in $\mathcal{E}_\Phi$ may be taken to be the equalizer in $\mathcal{E}$ of any pair of representing Φ-maps with the same domain $A \times U$. Similarly, the correspondence between morphisms $C \times A \times U \to B$ and morphisms $C \times U \to B^A$ in $\mathcal{E}$ (where U is any member of Φ) respects the equivalence relation, and so B^A remains an exponential in $\mathcal{E}_\Phi$. Monomorphisms in $\mathcal{E}_\Phi$ turn out to be equivalence classes of Φ-maps which include at least one monomorphism in $\mathcal{E}$, from which it is again easy to verify that Ω remains a subobject classifier in $\mathcal{E}_\Phi$.

For example, if $\mathcal{E} = \mathbf{Sh}(X)$, we may take Φ to be the filter of open neighbourhoods of some point $x \in X$, to obtain a topos $\mathbf{Sh}(X)_\Phi$ which we may think of as 'the germ of the topos of sheaves on X at the point x', in the sense that its objects are sheaves on X but its morphisms are what are generally known as *germs* at x of sheaf morphisms.

Again, if Φ is a filter (in the classical sense) on a set I, we may regard it as a filter of subobjects of 1 in $\mathbf{Set}/I$, and so form the quotient $(\mathbf{Set}/I)_\Phi$, which is commonly called a *filterpower* (or, if Φ is an ultrafilter, an *ultrapower*) of $\mathbf{Set}$. For an ultrafilter Φ, the topos $(\mathbf{Set}/I)_\Phi$ shares many of the properties of $\mathbf{Set}$; for example, it is Boolean (as are all filterquotients of Boolean toposes), and its terminal object $P_\Phi(1)$ is a generator and has just two isomorphism classes of subobjects. (We sometimes express the latter condition by saying that a topos is *2-valued*.) As we shall see later (in C1.4.10), these properties suffice to characterize $\mathbf{Set}$ (up to equivalence) amongst toposes which enjoy the non-elementary properties of cocompleteness and local smallness; however, if Φ is a non-principal ultrafilter then $(\mathbf{Set}/I)_\Phi$ is not cocomplete (see [32]), which demonstrates that some such non-elementary condition is essential for any characterization of $\mathbf{Set}$ amongst toposes.

Suggestions for further reading: Adelman & Johnstone [32], Dubuc & Kelly [306], Johnstone [507], Joyal [552], Lawvere [700–704, 712, 716], Mac Lane & Moerdijk [751, 752], Mulry [864].

A2.2 The Monadicity Theorem

The original definition of an elementary topos, as given by Lawvere and Tierney, included the requirement that $\mathcal{E}$ be cocartesian as well as cartesian. It was first shown by C. J. Mikkelsen [811] that this condition is redundant; subsequently R. Paré [927] gave a different proof of the same result, which established on the way the remarkable fact that if $\mathcal{E}$ is a topos then $\mathcal{E}^{\mathrm{op}}$ is monadic over $\mathcal{E}$. It is this proof which we describe in the present section.

Let $\mathcal{E}$ be a (weak) topos. We begin by noting that the assignment $A \mapsto PA$ can be made into the object-map of a functor $\mathcal{E}^{\mathrm{op}} \to \mathcal{E}$: if $f\colon A \to B$ is a morphism of $\mathcal{E}$, we define $Pf\colon PB \to PA$ to be the name of the relation $E_f \rightarrowtail PB \times A$

defined by the pullback square

$$\begin{tikzcd} E_f \arrow[r] \arrow[d, tail] & \in_B \arrow[d, tail] \\ PB \times A \arrow[r, "1 \times f"] & PB \times B \end{tikzcd} .$$

(It is straightforward to verify that this is functorial.) P may also be made into a covariant functor on the subcategory $\mathcal{E}_m$ of monomorphisms in $\mathcal{E}$: if $m: A \rightarrowtail B$ is monic, we define $\exists m: PA \to PB$ to be the name of the composite

$$\in_A \rightarrowtail PA \times A \overset{1 \times m}{\rightarrowtail} PA \times B .$$

(This notation is deliberately reminiscent of the functors $\exists_f$ defined for a regular category in Section A1.3; once we have seen that a topos is regular, we shall be able to extend the definition of $\exists f$ to an arbitrary morphism f of $\mathcal{E}$, and we shall see that $\exists f$ is simply the 'internal' version of $\exists_f$.)

Lemma 2.2.1 *The functor $P: \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$ has a left adjoint, namely $P: \mathcal{E} \to \mathcal{E}^{\mathrm{op}}$.*

Proof Unravelling the definitions, we have to set up a natural bijection between morphisms $A \to PB$ and morphisms $B \to PA$ in $\mathcal{E}$, or equivalently between (isomorphism classes of) subobjects of $A \times B$ and of $B \times A$. But the twist isomorphism $A \times B \to B \times A$ obviously induces such a bijection, which is natural in A and B. $\square$

We next introduce an important morphism from an object to its power object: for any object A of $\mathcal{E}$, let $\{\}: A \to PA$ be the name of the diagonal relation $(1_A, 1_A): A \rightarrowtail A \times A$. ($\{\}$ is pronounced 'singleton'; in the topos **Set** it is the mapping from a set to its power-set which sends an element a to the subset $\{a\}$.)

Lemma 2.2.2 *Let $f: A \to B$ be a morphism in a weak topos. Then*

(i) *The composite $\{\}f: A \to B \to PB$ names the relation $(1_A, f): A \rightarrowtail A \times B$.*

(ii) *The composite $Pf\{\}: B \to PB \to PA$ names the relation $(f, 1_A): A \rightarrowtail B \times A$.*

Proof (i) Consider the diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, tail, "{(1,f)}"] & B \arrow[r] \arrow[d, tail, "{(1,1)}"] & \in_B \arrow[d, tail] \\ A \times B \arrow[r, "f \times 1"] & B \times B \arrow[r, "\{\} \times 1"] & PB \times B \end{tikzcd}$$

in which both squares are pullbacks.

(ii) follows from (i) and the naturality of the bijection ($\Theta_{A,B}$, say) between morphisms $A \to PB$ and morphisms $B \to PA$, since $\Theta_{B,B}$ clearly maps $\{\}$ to itself. □

Corollary 2.2.3
(i) $\{\}\colon B \to PB$ *is monic for any* B.
(ii) *The functor* $P\colon \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$ *is conservative.*

Proof (i) Suppose $f, g\colon A \rightrightarrows B$ are such that $\{\}f = \{\}g$. Then $(1_A, f)$ and $(1_A, g)$ are isomorphic in $\mathrm{Sub}(A \times B)$; but this means that we have an isomorphism $i\colon A \to A$ such that $gi = f$ and $1_A i = 1_A$, whence $i = 1_A$ and $f = g$.

(ii) Since $\mathcal{E}$ (and hence $\mathcal{E}^{\mathrm{op}}$) is a balanced category by 1.6.2, it suffices to prove that P is faithful. Let $f, g\colon A \rightrightarrows B$ be such that $Pf = Pg$; then $Pf\{\} = Pg\{\}$, so we have $(f, 1_A) \cong (g, 1_A)$ in $\mathrm{Sub}(B \times A)$, and as before this forces $f = g$. □

The next result we require is a relation between the morphisms Pf and $\exists f$, often called the *Beck–Chevalley condition.*

Lemma 2.2.4 *Let*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, tail, "g"] & B \arrow[d, tail, "h"] \\ C \arrow[r, "k"] & D \end{tikzcd}$$

be a pullback square in a weak topos with g *and* h *monic. Then the square*

$$\begin{tikzcd} PB \arrow[r, "Pf"] \arrow[d, "\exists h"] & PA \arrow[d, "\exists g"] \\ PD \arrow[r, "Pk"] & PC \end{tikzcd}$$

commutes.

Proof Consider the diagram

$$\begin{tikzcd} E_f \arrow[r] \arrow[d, tail] & \in_B \arrow[d, tail] \\ PB \times A \arrow[r, "1 \times f"] \arrow[d, tail, "1 \times g"] & PB \times B \arrow[d, tail, "1 \times h"] \\ PB \times C \arrow[r, "1 \times k"] & PB \times D \end{tikzcd}$$

in which both squares are pullbacks: the upper one by definition, and the lower one because $PB \times (-)$ preserves pullbacks. It is easy to see that the left vertical composite, as a relation from PB to C, is named by both $\exists g \circ Pf$ and $Pk \circ \exists h \colon PB \to PC$. □

Corollary 2.2.5 *If $m \colon A \rightarrowtail B$ is monic, then $Pm \circ \exists m = 1_{PA}$.*

Proof Apply 2.2.4 to the pullback square

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, tail, "1"] & A \arrow[d, tail, "m"] \\ A \arrow[r, "m"] & B \end{tikzcd} .$$

□

We digress briefly to note, for future reference, the 'external' version of 2.2.5:

Lemma 2.2.6 *In a (weak) topos, any object of the form PA is injective. In particular, a topos has 'enough injectives'; i.e. for any object A there exists a monomorphism from A to an injective object.*

Proof Suppose given a pair of morphisms $m \colon B' \rightarrowtail B$, $f' \colon B' \to PA$ where m is monic. Let $R \rightarrowtail B' \times A$ be the relation named by f', and let $f \colon B \to PA$ name the composite

$$R \rightarrowtail B' \times A \xrightarrow{m \times 1} B \times A .$$

Then fm names the pullback of the latter subobject along $m \times 1$, which is isomorphic to $R \rightarrowtail B' \times A$ because m is monic; so $fm = f'$. The second assertion follows from the first and 2.2.3(i). □

Theorem 2.2.7 *For a weak topos $\mathcal{E}$, the functor $P \colon \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$ is monadic, i.e. it induces an equivalence between $\mathcal{E}^{\mathrm{op}}$ and the category of algebras for the monad on $\mathcal{E}$ induced by the adjunction of Lemma* 2.2.1.

Proof We shall show that $P \colon \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$ satisfies the three conditions of 1.1.2. The first two were verified in 2.2.1 and 2.2.3(ii) respectively, and $\mathcal{E}^{\mathrm{op}}$ certainly has coequalizers because $\mathcal{E}$ has equalizers. So it remains to prove that P sends

equalizers of coreflexive pairs in $\mathcal{E}$ to coequalizers. Let

$$E \xrightarrow{m} A \underset{g}{\overset{f}{\rightrightarrows}} B$$

be an equalizer diagram in $\mathcal{E}$ such that the pair (f, g) is coreflexive. Then by 1.2.10 the square

$$\begin{tikzcd} E \arrow[r, "m"] \arrow[d, "m"'] & A \arrow[d, "f"] \\ A \arrow[r, "g"] & B \end{tikzcd}$$

is a pullback. Applying 2.2.4 to it, we have $Pf \circ \exists g = \exists m \circ Pm$; but we also have $Pm \circ \exists m = 1_{PE}$ and $Pg \circ \exists g = 1_{PA}$ by 2.2.5, since m and g are monic, and $Pm \circ Pf = Pm \circ Pg$ since P is a functor. So the diagram

$$PB \underset{Pg}{\overset{Pf}{\rightrightarrows}} PA \xrightarrow{Pm} PE,$$

together with the morphisms $\exists g$ and $\exists m$, forms a split coequalizer system in $\mathcal{E}$; in particular it is a coequalizer diagram, as required. □

Remark 2.2.8 For the case $\mathcal{E} = \mathbf{Set}$, the monadicity result of 2.2.7 is well known. It follows from the Lindenbaum–Tarski representation theorem for complete atomic Boolean algebras that $\mathbf{Set}^{\mathrm{op}}$ is equivalent to the category of such algebras (equivalently, of completely distributive complete Boolean algebras, or of compact Hausdorff topological Boolean algebras; cf. [520, VI 1.16]), and it follows from the general results of [726] that such a category is monadic over **Set**. In a general topos $\mathcal{E}$, a power object PA is of course not a Boolean algebra but a Heyting algebra; but we shall see in Section C3.5 that there is a sense in which power objects may be regarded as 'complete atomic Heyting algebras' in $\mathcal{E}$, leading to a similar 'algebraic' description of $\mathcal{E}^{\mathrm{op}}$.

Corollary 2.2.9 *A (weak) topos is cocartesian.*

Proof Any category monadic over a cartesian category is cartesian, since the forgetful functor from a category of algebras creates limits. □

Corollary 2.2.10 *Let $F: \mathcal{E} \to \mathcal{F}$ be a logical functor between weak toposes. Then*

(i) *F is cocartesian.*

(ii) *F has a left adjoint iff it has a right adjoint.*

Proof Since F is logical, the diagram

$$\begin{tikzcd} \mathcal{E}^{\mathrm{op}} \arrow[r, "F"] \arrow[d, "P"] & \mathcal{F}^{\mathrm{op}} \arrow[d, "P"] \\ \mathcal{E} \arrow[r, "F"] & \mathcal{F} \end{tikzcd}$$

commutes up to natural isomorphism (it is easily seen that, for a general cartesian functor F, the canonical morphisms $\phi_A \colon F(PA) \to P(FA)$ defined in Section A2.1 form a natural transformation between the two composites). The first assertion is now immediate: $F \colon \mathcal{E}^{\mathrm{op}} \to \mathcal{F}^{\mathrm{op}}$ is cartesian because $F \colon \mathcal{E} \to \mathcal{F}$ is and $P \colon \mathcal{F}^{\mathrm{op}} \to \mathcal{F}$ creates limits. For the second, we use 1.1.3: the hypotheses of both parts of 1.1.3 are easily seen to be satisfied in the present case, but a left (resp. right) adjoint for $F \colon \mathcal{E}^{\mathrm{op}} \to \mathcal{F}^{\mathrm{op}}$ is the same thing as a right (resp. left) adjoint for $F \colon \mathcal{E} \to \mathcal{F}$. □

Remark 2.2.11 For future reference, it will be useful to write down the explicit construction of a right adjoint R for a logical functor F, given a left adjoint L for it: following the construction in [503], we find that for any object A of $\mathcal{F}$ we have an equalizer diagram

$$\begin{tikzcd} RA \arrow[r, tail] & PLPA \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & PLPPPA, \end{tikzcd}$$

where the morphisms f and g are obtained as follows: f is $PLP\eta_A$, where η is the unit of the adjunction of 2.2.1 (i.e. $\eta_A \colon A \to PPA$ names the relation $\in_A \rightarrowtail PA \times A \cong A \times PA$), and g corresponds under the adjunction of 2.2.1 to $\gamma_{PA} \colon LPPPA \to PPLPA$, where $\gamma \colon LPP \to PPL$ is the 'mate' under the adjunction $(L \dashv F)$ of the natural isomorphism $PPF \cong FPP$.

Suggestions for further reading: Mikkelsen [811], Paré [927].

A2.3 The Fundamental Theorem

The 'fundamental theorem of topos theory', so named by P. Freyd [371], is the assertion that the property of being a topos is stable under slicing. Whether it really deserves this name is perhaps open to debate (we shall see that it is not at all hard to prove), but there is no doubt that a great many important consequences flow from it, as we shall see in this section.

We begin with a lemma which is reminiscent of 1.5.2(ii) and 1.6.7 (indeed, 1.6.7 is a special case of it).

Lemma 2.3.1 *Let A be an object of a cartesian category $\mathcal{E}$ having a power object PA. Then, for any object B of $\mathcal{E}$, the object $B^*(PA)$, equipped with the relation $B^*(\in_A)$, is a power object for $B^*(A)$ in $\mathcal{E}/B$.*

Proof Let $f\colon C \to B$ be an object of $\mathcal{E}/B$. The square

$$\begin{tikzcd} C \times A \arrow[r, "f \times 1"] \arrow[d] & B \times A \arrow[d] \\ C \arrow[r, "f"] & B \end{tikzcd}$$

is a pullback, so $\Sigma_B(f \times B^*(A))$ is isomorphic to $C \times A$, and hence

$$\mathrm{Sub}_{\mathcal{E}/B}(f \times B^*(A)) \cong \mathrm{Sub}_{\mathcal{E}}(C \times A)\ .$$

Thus, given a relation from f to $B^*(A)$ in $\mathcal{E}/B$, we may regard it as a relation from C to A in $\mathcal{E}$; if $g\colon C \to PA$ is its name in $\mathcal{E}$, then $(g, f)\colon f \to B^*(PA)$ is the unique morphism naming it in $\mathcal{E}/B$. □

Note that, given the result which follows (and 2.3.7(iii) below), the isomorphism $\Sigma_B(f \times B^*(A)) \cong C \times A$, which we used in the above proof, is actually a special case of 1.5.8.

Theorem 2.3.2 *Let $\mathcal{E}$ be a weak topos, B an object of $\mathcal{E}$. Then $\mathcal{E}/B$ is a weak topos, and the pullback functor $B^*\colon \mathcal{E} \to \mathcal{E}/B$ is logical.*

Proof Given the first assertion, the second follows immediately from Lemma 2.3.1. To prove the first, we have to construct a power object for an arbitrary object $f\colon A \to B$ of $\mathcal{E}/B$, not just for one which happens to lie in the image of B^*. But, given f, the morphism $(1_A, f)\colon A \to A \times B$ is a monomorphism $f \rightarrowtail B^*(A)$ in $\mathcal{E}/B$ (in fact the unit of the adjunction $(\Sigma_B \dashv B^*)$ at the object f), and so we should expect the power object of f to be a subobject (in fact a retract, by 2.2.5) of the power object $B^*(PA)$ of $B^*(A)$.

Form the pullback

$$\begin{tikzcd} Q \arrow[rr] \arrow[d, tail, "{(h,k)}"] & & P_1A \arrow[d, tail] \\ PA \times B \arrow[r, "1 \times \{\}"] & PA \times PB \arrow[r, "1 \times Pf"] & PA \times PA \end{tikzcd}$$

where P_1A, the order-relation on PA, is defined as the equalizer of the first projection π_1 and the morphism $\wedge_A\colon PA \times PA \to PA$ which names the subobject $\pi_{13}^*(\in_A) \cap \pi_{23}^*(\in_A)$ of $PA \times PA \times A$ (cf. the definition of the order-relation on

Ω in the proof of 1.6.3). Now, given an object $g: C \to B$ of $\mathcal{E}/B$, morphisms $g \to k$ in $\mathcal{E}/B$ correspond to morphisms $l: C \to Q$ such that $kl = g$, or equivalently to morphisms $hl: C \to PA$ such that the pair $(hl, Pf\{\}g)$ factors through $P_1A \rightarrowtail PA \times PA$, or equivalently again to subobjects of $C \times A$ contained in the subobject named by $Pf\{\}g$. But, by the definition of $\{\}$, the latter composite names the pullback of the diagonal $B \rightarrowtail B \times B$ along $g \times f: C \times A \to B \times B$, i.e. the subobject $C \times_B A \rightarrowtail C \times A$. So we have a bijective correspondence between morphisms $g \to k$ and subobjects of $g \times f$ in $\mathcal{E}/B$, which is clearly natural in g (and hence induced by pulling back the subobject $\in_f \rightarrowtail k \times f$ which corresponds to the identity morphism $k \to k$); so k is a power object for f. □

Corollary 2.3.3 *For any morphism $f: A \to B$ in a weak topos $\mathcal{E}$, the functor $f^*: \mathcal{E}/B \to \mathcal{E}/A$ is logical, and has both left and right adjoints.*

Proof The first assertion follows from 2.3.2 and the isomorphism $\mathcal{E}/A \cong (\mathcal{E}/B)/f$. For the second, f^* always has a left adjoint by 1.2.8, so it has a right adjoint by 2.2.10(ii). □

Corollary 2.3.4 *A weak topos is locally cartesian closed.*

Proof This follows at once from 2.3.3 and 1.5.3. □

In particular, it follows from 2.3.4 that a weak topos is a topos, as defined in 2.1.1, so we can now drop the adjective 'weak'.

Corollary 2.3.5 *A topos is a Heyting category.*

Proof It is locally cartesian closed by 2.3.4 and cocartesian by 2.2.9, so this follows from 1.5.13(ii). □

In particular, a topos is a regular category; so we may now fulfil a promise made in the last section, and extend the definition of the covariant power-object functor to arbitrary morphisms of $\mathcal{E}$, rather than just monomorphisms. Given $f: A \to B$, we define $\exists f: PA \to PB$ to be the name of the relation which is the image of the composite

$$\in_A \rightarrowtail PA \times A \xrightarrow{1 \times f} PA \times B .$$

Since images are stable under pullback, it follows easily that if $g: C \to PA$ names a relation $R \rightarrowtail C \times A$, then $\exists f \circ g$ names the relation $\exists_{1 \times f}(R) \rightarrowtail C \times B$. From this in turn we deduce that the assignment $f \mapsto \exists f$ is functorial on $\mathcal{E}$. We note a few further properties of this functor:

Lemma 2.3.6

(i) $\{\}$ *is a natural transformation from the identity to the covariant power-object functor.*

(ii) *The Beck–Chevalley condition of Lemma 2.2.4 remains valid without the assumption that g and h are monic.*

(iii) *If* $f: A \to B$ *is epic, then* $\exists f \circ Pf = 1_{PB}$.

Proof (i) Given $f: A \to B$, we have to show that

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "\{\}"] & B \arrow[d, "\{\}"] \\ PA \arrow[r, "\exists f"] & PB \end{tikzcd}$$

commutes. By the remarks above, $\exists f \circ \{\}$ names the image of the composite

$$A \overset{(1,1)}{\rightarrowtail} A \times A \xrightarrow{1 \times f} A \times B;$$

but this composite is already monic, and by 2.2.2(i) it is also named by $\{\}f$.

(ii) The proof is essentially the same as that of 2.2.4; we have to replace the vertical composites in the diagram appearing in that proof by their images, but the resulting square is still a pullback by regularity.

(iii) If f is epic, then the top arrow in the pullback square

$$\begin{tikzcd} E_f \arrow[r] \arrow[d, tail] & \in_B \arrow[d, tail] \\ PB \times A \arrow[r, "1 \times f"] & PB \times B \end{tikzcd}$$

is also epic by regularity (note that, because a topos is balanced, all its epimorphisms are covers by the argument in the proof of 1.4.9). So the image of $E_f \rightarrowtail PB \times A \to PB \times B$ (which is the relation named by $\exists f \circ Pf$) is just $\in_B$. □

We have not yet proved, as promised in Section A2.1, that logical functors between toposes preserve the cartesian closed structure; but this can be done without much further effort. Let $F: \mathcal{E} \to \mathcal{F}$ be any functor between cartesian categories; then, for each object B of $\mathcal{E}$, F induces a functor $\mathcal{E}/B \to \mathcal{F}/FB$ (which we shall denote F/B) in an obvious way, and for each morphism $f: A \to B$ in $\mathcal{E}$ we have a diagram

$$\begin{tikzcd} \mathcal{E}/A \arrow[r, "\Sigma_f"] \arrow[d, "F/A"] & \mathcal{E}/B \arrow[r, "f^*"] \arrow[d, "F/B"] & \mathcal{E}/A \arrow[d, "F/A"] \\ \mathcal{F}/FA \arrow[r, "\Sigma_{Ff}"] & \mathcal{F}/FB \arrow[r, "(Ff)^*"] & \mathcal{F}/FA \end{tikzcd}$$

of which the left-hand square commutes by definition, and the right-hand square commutes provided F preserves pullbacks.

Proposition 2.3.7 *Let $F\colon \mathcal{E} \to \mathcal{F}$ be a logical functor between toposes.*

(i) *For each object B of $\mathcal{E}$, the functor $F/B\colon \mathcal{E}/B \to \mathcal{F}/FB$ is logical.*

(ii) *For each $f\colon A \to B$ in $\mathcal{E}$, the diagram*

$$\begin{tikzcd} \mathcal{E}/A \arrow[r, "\Pi_f"] \arrow[d, "F/A"] & \mathcal{E}/B \arrow[d, "F/B"] \\ \mathcal{F}/FA \arrow[r, "\Pi_{Ff}"] & \mathcal{F}/FB \end{tikzcd}$$

commutes up to isomorphism.

(iii) *F is a cartesian closed functor.*

Proof (i) This follows from the construction of power objects in $\mathcal{E}/B$ given in the proof of 2.3.2, since F preserves all the structure involved in this construction.

(ii) This similarly follows from the explicit construction of right adjoints for logical functors given at the end of Section A2.2: since the functors F/A and F/B are logical (i.e. commute up to isomorphism with power-object functors) and commute with the logical functor f^* and its left adjoint Σ_f, they preserve (in an appropriate sense) everything involved in the construction of the right adjoint Π_f.

(iii) The functor $(-)^B\colon \mathcal{E} \to \mathcal{E}$ may be factored as $\Pi_B \circ B^*$ (cf. the proof of 1.5.2(i)), so it follows from (ii) that the square

$$\begin{tikzcd} \mathcal{E} \arrow[r, "(-)^B"] \arrow[d, "F"] & \mathcal{E} \arrow[d, "F"] \\ \mathcal{F} \arrow[r, "(-)^{FB}"] & \mathcal{F} \end{tikzcd}$$

commutes up to isomorphism, i.e. F preserves exponentials. □

We have seen that pullback functors in a topos are logical, and that they have left (and right) adjoints. The latter property does not quite suffice to characterize them amongst logical functors between toposes, but there are several alternative hypotheses that we can add to obtain such a characterization. The next proposition collects together three of these; a fourth such condition will be found in C3.5.4(iii).

Proposition 2.3.8 *Let $F\colon \mathcal{E} \to \mathcal{F}$ be a logical functor between toposes having a left adjoint L. The following are equivalent:*

(i) *There is an object B of $\mathcal{E}$ and an equivalence $\mathcal{F} \simeq \mathcal{E}/B$ identifying F and L with B^* and Σ_B respectively.*

(ii) *L is faithful.*

(iii) *L preserves equalizers.*

(iv) *L preserves pullbacks.*

Proof (i) $\Rightarrow$ (ii), (iii) and (iv) because Σ_B is well known to have all the stated properties.

(ii) $\Rightarrow$ (i): Let $B = L(1)$. Since $\mathcal{F} \cong \mathcal{F}/1$, L 'lifts' to a functor $L'\colon \mathcal{F} \to \mathcal{E}/B$; and it is easily verified that L' has a right adjoint F', which sends $f\colon A \to B$ to the pullback of

$$\begin{tikzcd} & FA \arrow[d, "Ff"] \\ 1 \arrow[r, "\eta_1"] & FB \end{tikzcd}$$

(where η is the unit of $(L \dashv F)$). In other words, F' is the composite

$$\mathcal{E}/B \xrightarrow{F/B} \mathcal{F}/FB \xrightarrow{\eta_1^*} \mathcal{F},$$

and so it is logical by 2.3.7(i) and 2.3.3. Now L' preserves the terminal object by construction, and F' is cartesian closed by 2.3.7(iii), so applying 1.5.9(i) we deduce that the counit $L'F' \to 1_{\mathcal{E}/B}$ is an isomorphism. It now follows from one of the triangular identities that $L'\eta'\colon L' \to L'F'L'$ is an isomorphism, where η' is the unit of $(L' \dashv F')$; but L' clearly inherits faithfulness from L, and so reflects isomorphisms since $\mathcal{F}$ is balanced. Thus η' is an isomorphism, and L' and F' form the required equivalence of categories.

(iii) $\Rightarrow$ (ii): Since L preserves equalizers, its faithfulness (indeed, conservativeness, by 1.2.4) will follow provided we can show it preserves properness of subobjects. But the fact that F preserves Ω implies that we have a bijection, for any object A of $\mathcal{F}$, from $\mathrm{Sub}_{\mathcal{F}}(A)$ to $\mathrm{Sub}_{\mathcal{E}}(LA)$. We do not (yet) know that this bijection is simply the functor L applied to subobjects of A (but see 2.4.8 below); however, from the fact that the diagram

$$\begin{tikzcd} LA' \arrow[r] \arrow[d, tail] & L1 \arrow[r, "\epsilon_1"] \arrow[d, tail, "L\top"] & 1 \arrow[d, tail, "\top"] \\ LA \arrow[r, "L\phi"] & L\Omega \arrow[r, "\epsilon_\Omega"] & \Omega \end{tikzcd}$$

commutes (where ϕ is the classifying map of $A' \rightarrowtail A$, and ϵ is the counit of $(L \dashv F)$), we deduce that LA' is contained in the subobject of LA which corresponds to A', and in particular that it must be proper when A' is.

(iv) $\Rightarrow$ (iii) follows from 1.2.9. □

To show that the hypotheses of 2.3.8 are not vacuous, consider the case when $\mathcal{E} = \mathbf{Set}$ and $\mathcal{F} = [G, \mathbf{Set}]$ for a nontrivial group G, with $F: \mathcal{E} \to \mathcal{F}$ the functor which equips a set with the trivial G-action. Using 2.1.4, it is easy to see that F is logical; and it has a left adjoint L, which sends a G-set to the set of its G-orbits (and a right adjoint which sends a G-set to the set of its G-fixed points). But $\mathcal{F}$ is not equivalent to a slice of $\mathcal{E}$. By considering this example, one may verify that conditions (iii) and (iv) of 2.3.8 may not be weakened to 'L preserves equalizers of coreflexive pairs' or 'L preserves pullbacks of monomorphisms'; in fact, as we shall see in the next section, these conditions are satisfied by any functor left adjoint to a logical functor.

Before leaving this topic, we record

Scholium 2.3.9 *If a logical functor between toposes has a cartesian left adjoint, then it is an equivalence.*

Proof In the notation of 2.3.8, L preserves 1, so we have $\mathcal{F} \simeq \mathcal{E}/1 \cong \mathcal{E}$. □

Let $f: A \to B$ be a morphism in a topos $\mathcal{E}$. Just as the left adjoint $\exists_f: \mathrm{Sub}(A) \to \mathrm{Sub}(B)$ of f^* is 'internalized' by a morphism $\exists f: PA \to PB$ of $\mathcal{E}$, so the right adjoint $\forall_f$ of f^* (which exists by 2.3.5) corresponds to a morphism $\forall f: PA \to PB$. To define the latter, we simply take the name of the relation $\forall_{1 \times f}(\in_A) \rightarrowtail PA \times B$. Since the functors $\forall_g$ commute with pullback by 1.4.11, it follows easily that if $r: C \to PA$ names a relation $R \rightarrowtail C \times A$, then the composite $(\forall f)r$ is the name of $\forall_{1\times f}(R) \rightarrowtail C \times B$. From this, we may easily establish the analogue for $\forall$ of the Beck–Chevalley condition of 2.2.4 and 2.3.6(ii):

Lemma 2.3.10 *Let*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"] & B \arrow[d, "h"] \\ C \arrow[r, "k"] & D \end{tikzcd}$$

be a pullback square in a topos. Then the square

$$\begin{tikzcd} PB \arrow[r, "Pf"] \arrow[d, "\forall h"] & PA \arrow[d, "\forall g"] \\ PD \arrow[r, "Pk"] & PC \end{tikzcd}$$

commutes.

Proof By 1.4.11, each way round names the same subobject of $PB \times C$. $\square$

For future reference, we also note

Lemma 2.3.11 *Let $F: \mathcal{E} \to \mathcal{F}$ be a cartesian functor between toposes. Then F commutes with (external) universal quantification iff the diagram*

$$\begin{tikzcd} F(PA) \arrow[r, "F(\forall f)"] \arrow[d, "\phi_A"] & F(PB) \arrow[d, "\phi_B"] \\ P(FA) \arrow[r, "\forall(Ff)"] & P(FB) \end{tikzcd}$$

commutes for all $f: A \to B$ in $\mathcal{E}$. In particular, if F is a coherent functor, then it is a Heyting functor iff the above diagram commutes for all f.

Proof The two ways round the diagram name the subobjects $F(\forall_{1 \times f}(\in_A))$ and $\forall_{1 \times Ff}(F(\in_A))$ of $F(PA) \times FB$, so if F commutes with universal quantification then they are equal. The converse follows from composing the two ways round the diagram with the images under F of morphisms $1 \to PA$ in $\mathcal{E}$. $\square$

A2.4 Effectiveness, positivity and partial maps

We have seen that a topos is a Heyting category, and in particular regular; but we do not yet know that it is effective as a regular category. Our first task in this section is to prove that fact.

Let $(a, b): R \rightrightarrows A$ be an equivalence relation in a topos. The pair (a, b) is jointly monic, and so is named by a morphism $r: A \to PA$. In **Set**, r is the function which sends an element a of A to its equivalence class modulo R, so it is an obvious candidate for a morphism whose kernel-pair is (a, b). We now prove that this is true in general.

Proposition 2.4.1 *With the notation established above, (a, b) is a kernel-pair for r.*

Proof First we must show that $ra = rb: R \to PA$, or equivalently that the subobjects of $R \times A$ obtained by pulling back $(a, b): R \rightarrowtail A \times A$ along $a \times 1_A$ and $b \times 1_A$ are isomorphic. But if T denotes the 'object of R-related triples', i.e.

the pullback

$$\begin{tikzcd} T \arrow[r, "p"] \arrow[d, "q"'] & R \arrow[d, "b"] \\ R \arrow[r, "a"] & A \end{tikzcd}$$

(cf. 1.3.6(d)), then it follows easily from symmetry and transitivity of R that there are pullback squares

$$\begin{tikzcd} T \arrow[r, "t"] \arrow[d, tail, "{(p,bq)}"] & R \arrow[d, tail, "{(a,b)}"] \\ R \times A \arrow[r, "a \times 1"] & A \times A \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} T \arrow[r, "q"] \arrow[d, tail, "{(p,bq)}"] & R \arrow[d, tail, "{(a,b)}"] \\ R \times A \arrow[r, "b \times 1"] & A \times A \end{tikzcd}$$

(where t is the morphism expressing transitivity of (a, b)).

Now suppose given $(f, g)\colon B \rightrightarrows A$ with $rf = rg$. Then the subobjects $(f \times 1_A)^*(a, b)$ and $(g \times 1_A)^*(a, b)$ of $B \times A$ are isomorphic; pulling them back along $(1_B, g)\colon B \to B \times A$, we deduce that $(f, g)^*(a, b) \cong (g, g)^*(a, b)$ in $\mathrm{Sub}(B)$. But the latter is the whole of B, since (g, g) factors through the diagonal $A \rightarrowtail A \times A$ and (a, b) is reflexive; so $(f, g)^*(a, b)$ is the whole of B, i.e. (f, g) factors through (a, b). And the factorization is unique since (a, b) is jointly monic; so (a, b) has the universal property of a kernel-pair. □

For an alternative proof of 2.4.1, in a slightly more general setting, see 3.4.8(ii) below.

Note that we have not (explicitly) constructed a coequalizer for (a, b) in the proof of 2.4.1, since the morphism r is not epic (we shall see in D4.1.8(i) that there can never be an epimorphism $A \twoheadrightarrow PA$ in a nondegenerate topos). Of course, we may obtain a coequalizer for (a, b) by forming the image factorization of r; cf. 1.3.4 (and 3.4.8(iii) below).

It also seems worth remarking that 2.4.1 includes as a particular case the result, already proved in 2.2.3(i), that $\{\}: A \to PA$ is monic, since its kernel-pair is $(1_A, 1_A)$.

The fact that a topos is positive as a coherent category can be deduced from 1.5.14 and 1.6.2. However, we shall give a different proof of it, which establishes a result of independent interest along the way. Let $f\colon A \rightarrowtail B$, $g\colon A \to C$ be a pair of morphisms with common domain, the first of which is monic. Then $(f, g)\colon A \to B \times C$ is also monic, and so is named by a morphism $h\colon B \to PC$.

Lemma 2.4.2 *With the above notation, the square*

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, tail, "f"'] & C \arrow[d, tail, "\{\}"] \\ B \arrow[r, "h"] & PC \end{tikzcd}$$

is a pullback.

Proof The square commutes, because we have a pullback square

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, tail, "(1,g)"'] & A \arrow[d, tail, "(f,g)"] \\ A \times C \arrow[r, "f \times 1"] & B \times C \end{tikzcd},$$

and so hf names the subobject $(1, g)$ of $A \times C$, which is also named by $\{\}g$ (2.2.2(i)). Given any pair of morphisms $k\colon D \to B$, $l\colon D \to C$ with $hk = \{\}l$, we similarly see that there must be a pullback square

$$\begin{tikzcd} D \arrow[r] \arrow[d, tail, "(1,l)"'] & A \arrow[d, tail, "(f,g)"] \\ D \times C \arrow[r, "k \times 1"] & B \times C \end{tikzcd}.$$

The top edge of this square is the required factorization of (k, l) through (f, g) (and it is unique, because (f, g) is monic). □

Corollary 2.4.3 *Given a pushout square*

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, tail, "f"'] & C \arrow[d, "d"] \\ B \arrow[r, "e"] & D \end{tikzcd}$$

in a topos with f monic, the morphism d is monic and the square is also a pullback.

Proof Let $h: B \to PC$ be the name of (f, g), as in 2.4.2. By the universal property of pushouts, we have a factorization of $(h, \{\})$ through (e, d); since $\{\}$ is monic, this forces d to be monic, and the fact that the square of 2.4.2 is a pullback forces this one to be a pullback too. □

Corollary 2.4.4 *Coproducts in a topos are disjoint.*

Proof We have a pushout square

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & B \arrow[d] \\ A \arrow[r] & A+B \end{tikzcd} \quad ;$$

but the morphisms $0 \to A$ and $0 \to B$ are trivially monic, since 0 is a strict initial object by 1.4.1 (or by 1.5.12). So the result follows directly from 2.4.3. □

Combining 2.4.1 and 2.4.4, we have

Corollary 2.4.5 *A topos is a pretopos.* □

We remark that 2.4.4 applies also to infinite coproducts, to the extent that they exist in $\mathcal{E}$; also, a locally small topos is well-powered, since subobjects of A correspond to morphisms $A \to \Omega$. Thus *a locally small, cocomplete topos is an ∞-pretopos*. In fact there is a partial converse to this result, as we shall see in C2.2.8: an ∞-pretopos which has a separating set of objects is a topos.

We devote the rest of this section to a couple of mild digressions, which are included at least partly for historical reasons. The first concerns representability of partial maps: this condition originally formed one of the axioms for an elementary topos, when Lawvere and Tierney began their investigation of the subject. By the time the first published accounts of topos theory appeared, it had been realized that the condition could be deduced from a special case of itself (namely, the existence of a subobject classifier) in the presence of cartesian closedness, and so it no longer appeared as an axiom; but partial-map representers continued to play an important rôle in early accounts of the subject (e.g. [649], [1167], [504]) – in particular, in the proof of the Fundamental Theorem, where they were used in the construction of exponentials in $\mathcal{E}/B$. As we have seen in the preceding section, they are no longer needed for this purpose; but their existence and properties are still of interest.

A *partial map* from A to B in a cartesian category $\mathcal{E}$ is a special sort of relation from A to B, namely a subobject $R \rightarrowtail A \times B$ whose composite with the first projection $A \times B \to A$ is still monic. Equivalently, we may think of it as a pair of morphisms $(m: R \rightarrowtail A, f: R \to B)$ of which the first is monic. We shall write '$(m, f): A \rightharpoonup B$' (or sometimes simply '$f: A \rightharpoonup B$', suppressing the name of the monomorphism) for '(m, f) is a partial map from A to B'.

Partial maps can be composed: given $(m, f): A \rightharpoonup B$ and $(n, g): B \rightharpoonup C$, we form the pullback

$$\begin{tikzcd}
T \arrow[r] \arrow[d, tail] & S \arrow[r, "g"] \arrow[d, tail, "n"] & C \\
R \arrow[r, "f"] \arrow[d, tail, "m"] & B & \\
A & &
\end{tikzcd} .$$

This composition is associative up to (coherent) isomorphism, and so (provided we identify isomorphic relations) the objects and partial maps of $\mathcal{E}$ form a category $\mathbf{Part}(\mathcal{E})$. There is a conservative functor $\mathcal{E} \to \mathbf{Part}(\mathcal{E})$ sending an object A to itself and a morphism $f: A \to B$ to $(1_A, f): A \rightharpoonup B$.

We say that $\mathcal{E}$ has *representable partial maps* if this functor has a right adjoint: i.e. if, for each object A of $\mathcal{E}$, we are given an object $\tilde{A}$ and a partial map $(i_A, g_A): \tilde{A} \rightharpoonup A$ such that every partial map $B \rightharpoonup A$ factors uniquely through (i_A, g_A) by a morphism $B \to \tilde{A}$ in $\mathcal{E}$. This definition can be simplified: by an easy generalization of the argument used to prove Lemma 1.6.1, it can be shown that the second component g_A of such a universal partial map is necessarily an isomorphism, and so we may as well take it to be the identity on A. Thus we define a *partial-map representer* for A to be a monomorphism $i_A: A \rightarrowtail \tilde{A}$ such that, given a partial map $(m, f): B \rightharpoonup A$, there is a unique morphism $\tilde{f}: B \to \tilde{A}$ making the square

$$\begin{tikzcd}
R \arrow[r, "f"] \arrow[d, tail, "m"] & A \arrow[d, tail, "i_A"] \\
B \arrow[r, "\tilde{f}"] & \tilde{A}
\end{tikzcd}$$

a pullback.

A partial map $B \rightharpoonup 1$ is (essentially) the same thing as a subobject of B, so a partial-map representer for 1 is just another way of describing a subobject classifier. Moreover, the proof of Lemma 1.6.5 immediately generalizes to yield

Lemma 2.4.6 *A Boolean pretopos has representable partial maps.*

Proof Define $\tilde{A} = A \amalg 1$, and let i_A be the first coprojection. Given a partial map $(m, f): B \rightharpoonup A$, let m' be a complement for m in $\mathrm{Sub}(B)$, and define $\tilde{f}$ as

the unique morphism making

$$\begin{tikzcd} R \arrow[r, tail, "m"] \arrow[d, "f"] & B \arrow[d, "\tilde{f}"] & R' \arrow[l, tail, "m'"'] \arrow[d] \\ A \arrow[r, tail, "i"] & \tilde{A} & 1 \arrow[l, tail, "j"'] \end{tikzcd}$$

commute (where j is the second coprojection). The remaining details are as in 1.6.5. □

In a topos, we have already seen (in Lemma 2.4.2) that if we take $\tilde{A} = PA$ and $i_A = \{\}$, we get the 'existence' but not the 'uniqueness' part of the definition of a partial-map representer. The trouble is that PA is 'too big' (in **Set**, if A is an n-element set, PA has 2^n elements, whereas by 2.4.6 we want something with $n+1$ elements), and so we need to cut it down somehow.

Proposition 2.4.7 *A topos has representable partial maps.*

Proof Given an object A of $\mathcal{E} \cong \mathcal{E}/1$, let us write $\tau\colon \tilde{A} \to \Omega$ for the object $\Pi_\top(A)$ of $\mathcal{E}/\Omega$. We note that since $\top$ is monic, the unit of the adjunction ($\Sigma_\top \dashv \top^*$) is an isomorphism, and hence so is the counit of ($\top^* \dashv \Pi_\top$); in particular, $\top^*(\tau) \cong A$, and so we can write the subobject of $\tilde{A}$ classified by τ as $i_A : A \rightarrowtail \tilde{A}$. Given any partial map

$$\begin{tikzcd} B' \arrow[r, "f"] \arrow[d, tail, "m"] & A \\ B & \end{tikzcd} ,$$

let $\chi\colon B \to \Omega$ be the classifying map of m; then the morphism $f\colon B' = \top^*(\chi) \to A$ transposes to yield a morphism $\tilde{f}: \chi \to \tau$ in $\mathcal{E}/\Omega$ such that $\top^*(\tilde{f}) = f$, i.e. the square

$$\begin{tikzcd} B' \arrow[r, "f"] \arrow[d, tail, "m"] & A \arrow[d, tail, "i_A"] \\ B \arrow[r, "\tilde{f}"] & \tilde{A} \end{tikzcd}$$

is a pullback. Conversely, if this square is a pullback, then we have $\tau\tilde{f} = \chi$ by uniqueness of classifying maps, and $\tilde{f}$ must be the transpose of f. □

As an application of partial-map representers, we prove a result promised after 2.3.8:

Lemma 2.4.8 *Let $F\colon \mathcal{E} \to \mathcal{F}$ be a logical functor between toposes having a left adjoint L. Then L preserves monomorphisms; and in fact, for every object A of $\mathcal{F}$, the assignment $m \mapsto Lm$ is a bijection from (isomorphism classes of) subobjects of A in $\mathcal{F}$ to subobjects of LA in $\mathcal{E}$.*

Proof We clearly have a bijection between subobjects of A classified by morphisms $\phi\colon A \to \Omega_{\mathcal{F}} \cong F\Omega_{\mathcal{E}}$ and subobjects of LA classified by morphisms $\overline{\phi}\colon LA \to \Omega_{\mathcal{E}}$; the problem is to show that it is induced by the action of L on monomorphisms. For the moment, let us write $\lambda A' \rightarrowtail LA$ for the subobject corresponding to $m\colon A' \rightarrowtail A$. Since the diagram

$$\begin{tikzcd}
LA' \arrow[r] \arrow[d, "Lm"] & L1 \arrow[r, "\epsilon_1"] \arrow[d, "L\top"] & 1 \arrow[d, "\top"] \\
LA \arrow[r, "L\phi"] & L\Omega_{\mathcal{F}} \arrow[r, "\epsilon_\Omega"] & \Omega_{\mathcal{E}}
\end{tikzcd}$$

is readily seen to commute (where ϵ is the counit of $(L \dashv F)$), we have a canonical morphism $LA' \to \lambda A'$; to show that this is an isomorphism, we shall demonstrate that $\lambda A'$ has the same universal property as LA'.

For any object B of $\mathcal{E}$, morphisms $A' \to FB$ can be regarded as partial maps $A \rightharpoonup FB$, and so correspond to morphisms $f\colon A \to \widetilde{FB}$ satisfying $\delta f = \phi$, where $\delta\colon \widetilde{FB} \to \Omega_{\mathcal{F}}$ classifies $i_{FB}\colon FB \rightarrowtail \widetilde{FB}$. But since F is logical, it commutes with everything involved in the construction of 2.4.7; so we have an isomorphism $\widetilde{FB} \cong F(\tilde{B})$, which identifies δ with $F(\delta')$ where $\delta'\colon \tilde{B} \to \Omega_{\mathcal{E}}$ classifies i_B. Thus morphisms $A' \to FB$ correspond to morphisms $\overline{f}\colon LA \to \tilde{B}$ satisfying $\delta'\overline{f} = \overline{\phi}$, and these in turn correspond to morphisms $\lambda A' \to B$, as required. □

We note in passing that, since L induces isomorphisms of subobject lattices, it preserves intersections of subobjects (i.e. pullbacks of monomorphisms), and hence also equalizers of coreflexive pairs, as we claimed in the last section.

We have already observed that $\tilde{1}$ is another name for Ω. The following identification is also useful:

Lemma 2.4.9 *In a cartesian category with subobject classifier Ω, the partial-map representer $\tilde{\Omega}$ (exists and) is isomorphic to the order-relation Ω_1 defined in the proof of* 1.6.3.

Proof By definition, a partial map $A \rightharpoonup \Omega$ corresponds to a subobject $A' \rightarrowtail A$ (the domain of the partial map) and a subobject $A'' \rightarrowtail A'$ – equivalently, to a pair of subobjects $(A'' \rightarrowtail A, A' \rightarrowtail A)$ of which the first is contained in the second. But such a pair is classified by a morphism $A \to \Omega_1$. More explicitly,

let $i_\Omega\colon \Omega \to \Omega_1$ classify the pair of subobjects $(\top\colon 1 \rightarrowtail \Omega, 1_\Omega\colon \Omega \rightarrowtail \Omega)$; then i_Ω is monic, since its composite with the first projection is the identity on Ω (the classifying map of $\top$), and given a partial map $(A' \rightarrowtail A, f\colon A' \to \Omega)$ we let $\tilde{f}\colon A \to \Omega_1$ classify the pair (A'', A') where A'' is the subobject of A' classified by f. It is then straightforward to verify that $\tilde{f}$ is the unique morphism making

$$\begin{tikzcd} A' \arrow[r, "f"] \arrow[d, tail] & \Omega \arrow[d, tail, "i_\Omega"] \\ A \arrow[r, "\tilde{f}"] & \Omega_1 \end{tikzcd}$$

a pullback. □

Remark 2.4.10 There are a number of alternative constructions of $\tilde{A}$ which are sometimes of use. We mention two of them here, both of which are based on the idea that $\tilde{A}$ may be identified with the 'object of subobjects of A having at most one element'.

(a) Let $s\colon PA \to PA$ be the name of $(\{\}, 1_A)\colon A \rightarrowtail PA \times A$, and define $\tilde{A} \rightarrowtail PA$ to be the equalizer of s and 1_{PA}. (In **Set**, s is the map which sends each singleton subset of A to itself, and everything else to the empty set.) First we note that there is a pullback

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, tail, "{(1,1)}"] & A \arrow[d, tail, "{(\{\},1)}"] \\ A \times A \arrow[r, "\{\} \times 1"] & PA \times A \end{tikzcd}$$

because $\{\}$ is monic, and so $s\{\} = \{\}$; hence $\{\}$ factors through $\tilde{A} \rightarrowtail PA$ by a monomorphism $i_A\colon A \rightarrowtail \tilde{A}$. More generally, if $h\colon B \to PA$ names a partial map $(m, f)\colon B \rightharpoondown A$, then it follows immediately from 2.4.2 that the square

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, tail, "{(m,f)}"] & A \arrow[d, tail, "{(\{\},1)}"] \\ B \times A \arrow[r, "h \times 1"] & PA \times A \end{tikzcd}$$

is a pullback; so $sh = h$, and h factors through $\tilde{A} \rightarrowtail PA$ by a morphism $\tilde{f}\colon B \to \tilde{A}$. And the fact that the square in the statement of 2.4.2 is a pullback

implies that the square

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, tail, "m"] & A \arrow[d, tail, "i_A"] \\ B \arrow[r, "\tilde{f}"] & \tilde{A} \end{tikzcd}$$

is a pullback too.

It remains to show the uniqueness of $\tilde{f}$. Let $g\colon B \to \tilde{A}$ be another morphism making the last square above a pullback, and write g' for the composite of g with $\tilde{A} \rightarrowtail PA$. Then it is easy to see that the second square above, with h replaced by g', must also be a pullback, and so sg' names the subobject $(m, f)\colon R \rightarrowtail B \times A$. But $sg' = g'$ since g' factors through $\tilde{A} \rightarrowtail PA$, and so g' names this subobject; hence $g' = h$, and so $g = \tilde{f}$ since $\tilde{A} \rightarrowtail PA$ is monic.

We remark that, by similar methods, the square

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, tail, "(\{\}, 1)"] & A \arrow[d, tail, "(\{\}, 1)"] \\ PA \times A \arrow[r, "s \times 1"] & PA \times A \end{tikzcd}$$

can be shown to be a pullback; hence s is idempotent (i.e. $s^2 = s$), and we could equivalently have defined $\tilde{A}$ as the image of s. In particular, $\tilde{A}$ is a retract of PA, and is thus injective (this could have been proved directly from the definition of a partial-map representer, by methods similar to those of 2.2.6).

(*b*) For the second alternative construction, let $q\colon PA \to P(A \times A)$ name the relation

$$\pi_{12}^*(\in_A) \cap \pi_{13}^*(\in_A) \rightarrowtail PA \times A \times A$$

(in **Set**, q is the map which sends a subset A' of A to $A' \times A' \subseteq A \times A$), and let $d\colon 1 \to P(A \times A)$ name the diagonal map $A \rightarrowtail A \times A \cong 1 \times A \times A$. Form the pullback

$$\begin{tikzcd} \tilde{A} \arrow[r] \arrow[d, tail] & P_1(A \times A) \arrow[d, tail] \\ PA \cong PA \times 1 \arrow[r, "q \times d"] & P(A \times A) \times P(A \times A) \end{tikzcd}$$

where $P_1(A \times A)$, as before, denotes the order-relation on $P(A \times A)$. (We may think of $\tilde{A}$ as 'the object of subobjects of A whose squares are contained in the diagonal'.

To show that this definition has the required universal property, let $h\colon B \to PA$ name a relation $(g, f)\colon R \rightarrowtail B \times A$. We shall show that h factors through $\tilde{A} \rightarrowtail PA$ iff (g, f) is a partial map (i.e. iff g is monic).

But, by definition, h factors through $\tilde{A} \rightarrowtail PA$ iff $\pi_{12}^*(R) \cap \pi_{13}^*(R) \rightarrowtail B \times A \times A$ is contained in the subobject $1 \times \Delta\colon B \times A \rightarrowtail B \times A \times A$; equivalently, iff the top left square in the diagram

$$\begin{tikzcd}
R \arrow[r, tail] \arrow[d, tail] & \pi_{12}^*(R) \arrow[r] \arrow[d, tail] & R \arrow[d, tail, "{(g,f)}"] \\
\pi_{13}^*(R) \arrow[r, tail] \arrow[d] & B \times A \times A \arrow[r, "\pi_{12}"] \arrow[d, "\pi_{13}"] & B \times A \arrow[d, "\pi_1"] \\
R \arrow[r, tail, "{(g,f)}"] & B \times A \arrow[r, "\pi_1"] & B
\end{tikzcd}$$

(where the arrows $R \rightarrowtail \pi_{12}^*(R)$ and $R \rightarrowtail \pi_{13}^*(R)$ are the splittings of the projections induced by $1 \times \Delta\colon B \times A \rightarrowtail B \times A \times A$) is a pullback. And the other three squares in this diagram are pullbacks; so this happens iff the outside is a pullback, i.e. iff g is monic.

Our second (and much shorter!) digression concerns the construction of exponentials from power objects: we have seen that a weak topos must have exponentials, but we have not given an explicit procedure for constructing them. (Such a construction could in theory be extracted from 2.2.11 and 1.5.2, but it would be too complicated to be of much use.) The first proof of this result (due to A. Kock [636]) did yield an explicit construction of exponentials from power objects, which is based on the idea that in **Set** the exponential B^A may be identified with a subset of $P(A \times B)$, namely the set of subsets of $A \times B$ which are the graphs of functions $A \to B$. Kock's construction produces, for each pair of objects (A, B) in a (weak) topos, a subobject of $P(A \times B)$ which we may think of as 'the object of graphs of morphisms $A \to B$'; it is sometimes useful to have such an explicit construction, and so we give it here.

First we let $r\colon P(A \times B) \times A \to PB$ name the relation

$$\in_{(A \times B)} \rightarrowtail P(A \times B) \times (A \times B) \cong (P(A \times B) \times A) \times B,$$

and then we form the pullback

$$\begin{tikzcd}
Q \arrow[r] \arrow[d, tail] & B \arrow[d, tail, "\{\}"] \\
P(A \times B) \times A \arrow[r, "r"] & PB
\end{tikzcd} .$$

(In **Set**, Q is the set of pairs (R, a) such that a is R-related to exactly one element of B.) Now let $q\colon P(A \times B) \to PA$ be the name of the relation Q, and form the pullback

$$\begin{tikzcd} E \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top_A"] \\ P(A \times B) \arrow[r, "q"] & PA \end{tikzcd}$$

where $\top_A$ names the top element of $\mathrm{Sub}(1 \times A)$.

Proposition 2.4.11 *The object E constructed above has the universal property of an exponential B^A.*

Proof It suffices to show that a morphism $h\colon C \to P(A \times B)$ factors through $E \rightarrowtail P(A \times B)$ iff the relation $(m, n, p)\colon D \rightarrowtail C \times A \times B$ which it names has the property that $(m, n)\colon D \to C \times A$ is an isomorphism, i.e. iff it is the graph of a morphism $C \times A \to B$. But such a factorization exists iff qh names the top element of $\mathrm{Sub}(C \times A)$, which is equivalent to saying that $h \times 1_A\colon C \times A \to P(A \times B) \times A$ factors through $Q \rightarrowtail P(A \times B) \times A$, and this in turn is equivalent to saying that $r(h \times 1_A)$ factors through $\{\}\colon B \rightarrowtail PB$. But $r(h \times 1_A)$ names the subobject $D \rightarrowtail C \times A \times B$, considered as a relation from $C \times A$ to B; so this says that there is a (unique) morphism $f\colon C \times A \to B$ making

$$\begin{tikzcd} D \arrow[r] \arrow[d, tail, "{(m,n,p)}"] & B \arrow[d, "\Delta"] \\ C \times A \times B \arrow[r, "f \times 1"] & B \times B \end{tikzcd}$$

a pullback, i.e. that (m, n, p) is (isomorphic to) the graph of f. □

In passing, we remark that 2.4.11 provides an alternative proof of 2.3.7(iii): for a logical functor preserves all the structure used in the construction of the object E.

Suggestions for further reading: Kock & Mikkelsen [636], Kock & Wraith [649].

A2.5 Natural number objects

The axioms for a topos, as set out in Section A2.1, permit one to carry out a great many mathematical constructions 'inside' a topos, as we shall see later on in this book. However, lots of constructions in mathematics (in particular, those which involve the formation of recursively defined collections of entities) require

some form of axiom of infinity, and the topos axioms do not imply any such axiom (as is shown by the fact that $\mathbf{Set}_f$ is a model for them). So, for many purposes, we shall need to add such an axiom to the topos axioms.

The neatest categorical formulation of the axiom of infinity was found by F. W. Lawvere [695]:

Definition 2.5.1 Let $\mathcal{C}$ be a category with a terminal object 1. A *natural number object* in $\mathcal{C}$ is an object N equipped with morphisms $o\colon 1 \to N$ and $s\colon N \to N$ such that, given any object A with morphisms $a\colon 1 \to A$ and $t\colon A \to A$, there is a unique $f\colon N \to A$ such that

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "a"'] & N \arrow[r, "s"] \arrow[d, "f"] & N \arrow[d, "f"] \\ & A \arrow[r, "t"] & A \end{tikzcd}$$

commutes.

The definition says that (N, o, s) is an initial object in the category of $\mathbb{T}$-algebras in $\mathcal{C}$, where $\mathbb{T}$ is the finitary algebraic theory freely generated by one nullary and one unary operation. (Later we shall see that, in a topos, the existence of a natural number object implies the existence of free algebras for arbitrary finitely-presented algebraic theories.) It is thus immediate that a natural number object, if it exists, is unique up to canonical isomorphism. (Note that we shall frequently abuse notation by saying 'N is a natural number object', without mentioning the morphisms o and s. It should also be mentioned that many authors write 'natural numbers object' rather than 'natural number object'.)

In **Set**, the usual set $\mathbb{N}$ of (Cantor/von Neumann) natural numbers serves as a natural number object, with o and s taken to be 'zero' and 'successor'. The definition is a particular case of the assertion that functions may be defined by primitive recursion, but the general case follows from it:

Proposition 2.5.2 *Let (N, o, s) be a natural number object in a cartesian closed category. Then, given objects A and B equipped with morphisms $g\colon A \to B$ and $h\colon A \times N \times B \to B$, there exists a unique $f\colon A \times N \to B$ such that the diagram*

$$\begin{tikzcd} A \times 1 \arrow[r, "1 \times o"] \arrow[d, equal] & A \times N \arrow[d, "f"] & A \times N \arrow[l, "1 \times s"'] \arrow[d, "{(1, f)}"] \\ A \arrow[r, "g"] & B & A \times N \times B \arrow[l, "h"'] \end{tikzcd}$$

commutes.

Intuitively, f is the function defined recursively by $f(a,0) = g(a)$ and

$$f(a, n+1) = h(a, n, f(a,n)) .$$

Proof First we consider the case $A = 1$. By definition, we have a unique $f' \colon N \to N \times B$ such that

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "{(o,g)}"'] & N \arrow[r, "s"] \arrow[d, "f'"] & N \arrow[d, "f'"] \\ & N \times B \arrow[r, "{(s\pi_1, h)}"] & N \times B \end{tikzcd}$$

commutes; but the composite $\pi_1 f' \colon N \to N$ must be the identity, because it commutes with both o and s. So f' is the graph of a morphism $f = \pi_2 f' \colon N \to B$, which is clearly the unique morphism with the required properties.

The general case is reduced to this one by cartesian closedness. Let $\overline{g} \colon 1 \to B^A$ be the exponential transpose of g, and $\overline{h} \colon N \times B^A \to B^A$ the transpose of

$$N \times B^A \times A \xrightarrow{(\pi_3, \pi_1, \mathrm{ev} \circ \pi_{23})} A \times N \times B \xrightarrow{h} B .$$

Then it is easy to see that $f \colon A \times N \to B$ makes the diagram in the statement commute iff its transpose $\overline{f} \colon N \to B^A$ makes

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "\overline{g}"'] & N \arrow[d, "\overline{f}"] & N \arrow[l, "s"'] \arrow[d, "{(1,\overline{f})}"] \\ & B^A & N \times B^A \arrow[l, "\overline{h}"] \end{tikzcd}$$

commute. □

Remark 2.5.3 We note that the first half of the proof of 2.5.2 made no use of cartesian closedness. If we wish to study natural number objects in cartesian categories which are not cartesian closed, it makes sense to adopt a stronger form of the definition: (N, o, s) is a natural number object if, given any diagram $(A \to B \to B)$, there exists a unique morphism $f \colon A \times N \to B$ making the

diagram

$$\begin{tikzcd}
A \times 1 \arrow[r, "1_A \times o"] \arrow[d, "\pi_1"] & A \times N \arrow[r, "1_A \times s"] \arrow[d, "f"] & A \times N \arrow[d, "f"] \\
A \arrow[r] & B \arrow[r] & B
\end{tikzcd}$$

commute. (An equivalent condition would be to demand that N is a natural number object (in the sense of 2.5.1) 'stably under slicing', i.e. that A^*N is a natural number object in $\mathcal{C}/A$ for any A.) With this modification, the result of 2.5.2 remains true for natural number objects in arbitrary cartesian categories. However, we shall be almost exclusively concerned with natural number objects in toposes (or in quasitoposes; cf. Section A2.6 below); so we have chosen to adopt the simpler form of the definition given in 2.5.1.

By taking $B = N$ and A to be a finite power of N in 2.5.2, we deduce that every primitive recursive function of k variables is 'realized' as a morphism $N^k \to N$, in any cartesian closed category with a natural number object.

Examples 2.5.4 We define *addition*, *multiplication* and *exponentiation* on N recursively by the diagrams

$$\begin{tikzcd}
N \arrow[r, "1 \times o"] \arrow[dr, "1"'] & N \times N \arrow[d, "+"] & N \times N \arrow[l, "1 \times s"'] \arrow[d, "{(1,+)}"] \\
& N & N \times N \times N \arrow[l, "s\pi_3"']
\end{tikzcd}$$

(i.e. $m + 0 = m,\ m + (n+1) = (m+n)+1$),

$$\begin{tikzcd}
N \arrow[r, "1 \times o"] \arrow[d] & N \times N \arrow[d, "\times"] & N \times N \arrow[l, "1 \times s"'] \arrow[d, "{(1,\times)}"] \\
1 \arrow[r, "o"] & N & N \times N \times N \arrow[l, "{+(\pi_3,\pi_1)}"']
\end{tikzcd}$$

(i.e. $m \times 0 = 0,\ m \times (n+1) = (m \times n) + m$), and

$$\begin{tikzcd}
N \arrow[r, "1 \times o"] \arrow[d] & N \times N \arrow[d, "{[\leftarrow]}"] & N \times N \arrow[l, "1 \times s"'] \arrow[d, "{(1,[\leftarrow])}"] \\
1 \arrow[r, "so"] & N & N \times N \times N \arrow[l, "{\times(\pi_3,\pi_1)}"']
\end{tikzcd}$$

(i.e. $[m \leftarrow 0] = 1$, $[m \leftarrow (n+1)] = [m \leftarrow n] \times m$ – we use the notation $[m \leftarrow n]$ rather than the more familiar m^n in order to avoid confusion with exponentials within our category). Moreover, the usual inductive proofs translate, via the uniqueness clause of 2.5.2, into proofs that these operations satisfy the diagrammatic forms of the usual (associative, commutative, distributive) laws of arithmetic.

Two further morphisms which we shall require are the *predecessor map* $p\colon N \to N$ defined by

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "o"'] & N \arrow[d, "p"] & N \arrow[l, "s"'] \arrow[d, "{(1,p)}"] \\ & N & N \times N \arrow[l, "\pi_1"] \end{tikzcd}$$

and the *truncated subtraction* $\dot{-}\colon N \times N \to N$ defined by

$$\begin{tikzcd} N \arrow[r, "1 \times o"] \arrow[dr, "1"'] & N \times N \arrow[d, "\dot{-}"] & N \times N \arrow[l, "1 \times s"'] \arrow[d, "{(1,\dot{-})}"] \\ & N & N \times N \times N \arrow[l, "p\pi_3"] \end{tikzcd}$$

Note that the definition of p requires only the first half of the proof of 2.5.2, so this morphism exists in any category with finite products and a natural number object – in particular, $s\colon N \to N$ is a (split) monomorphism in any such category. This is essentially the fourth of Peano's five postulates characterizing the natural numbers in **Set** (the first two are just the existence of the morphisms o and s).

We now turn to a different characterization of natural number objects in toposes, due to P. Freyd [371].

Lemma 2.5.5 *Let (N, o, s) be a natural number object in a category $\mathbf{C}$.*

(i) *If $\mathbf{C}$ has either binary coproducts or binary products, then*

$$1 \xrightarrow{\ o\ } N \xleftarrow{\ s\ } N$$

is a coproduct diagram.

(ii) *For any $\mathbf{C}$,*

$$\begin{tikzcd} N \arrow[r, shift left, "1"] \arrow[r, shift right, "s"'] & N \arrow[r] & 1 \end{tikzcd}$$

is a coequalizer diagram.

Proof (i) If $\mathcal{C}$ has binary coproducts, then a natural number object in $\mathcal{C}$ is the same thing as an initial object in the category of T-algebras, where $T: \mathcal{C} \to \mathcal{C}$ is the endofunctor $1 + (-)$. So by 1.1.4 the morphism $1 + N \to N$ induced by o and s is an isomorphism.

In the case when $\mathcal{C}$ has binary products, suppose given an object A and morphisms $f: 1 \to A$, $g: N \to A$. Consider the unique morphism $(h, k): N \to N \times A$ making

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "(o,f)"'] & N \arrow[r, "s"] \arrow[d, "(h,k)"] & N \arrow[d, "(h,k)"] \\ & N \times A \arrow[r, "(s,g)\pi_1"] & N \times A \end{tikzcd}$$

commute: we have $ho = o$ and $hs = sh$, so $h = 1_N$ by uniqueness, and hence k is a factorization of (f, g) through (o, s). And it is unique, by the uniqueness of recursively defined morphisms.

(ii) If $f: N \to A$ is any morphism satisfying $fs = f$, then f is equal to the composite

$$N \longrightarrow 1 \xrightarrow{o} N \xrightarrow{f} A,$$

because each of them makes the diagram

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "fo"'] & N \arrow[r, "s"] \arrow[d] & N \arrow[d] \\ & A \arrow[r, "1"] & A \end{tikzcd}$$

commute. So f factors (uniquely, since $N \to 1$ is split epic) through $N \to 1$. □

Freyd's characterization asserts that in a topos the converse of Lemma 2.5.5 is true: if N is any object equipped with morphisms o and s such that the diagrams in the statement of 2.5.5 are colimits, then it is a natural number object. We postpone the proof of this result until Section D5.1, because it requires the development of the internal logic of a topos; nevertheless, we make use of it in part (ii) of the next result.

Lemma 2.5.6 *Let $\mathcal{C}$ and $\mathcal{D}$ be categories with terminal objects, and $F: \mathcal{C} \to \mathcal{D}$ a functor preserving the terminal object. Suppose $\mathcal{C}$ has a natural number object (N, o, s). If either* (i) *F has a right adjoint, or* (ii) *$\mathcal{C}$ has finite products or coproducts, $\mathcal{D}$ is a topos and F is cocartesian, then (FN, Fo, Fs) is a natural number object in $\mathcal{D}$.*

Proof (i) follows directly from the definition: given $a: 1 \to A$ and $t: A \to A$ in $\mathcal{D}$, we have

$$1 \cong R1 \xrightarrow{Ra} RA \xrightarrow{Rt} RA$$

in $\mathcal{C}$, where R is the right adjoint of F; so we get a unique morphism $f: N \to RA$ making the appropriate diagrams commute, and its transpose $FN \to A$ is the morphism we seek.

(ii) (N, o, s) satisfies the colimit diagrams of 2.5.5, and F preserves them; so the result follows from D5.1.2. □

The hypothesis that $\mathcal{D}$ is a topos in 2.5.6(ii) cannot be weakened; see C4.2.4.

Using 2.5.6, we may give many more examples of natural number objects. For example, $[\mathcal{C}, \mathbf{Set}]$ has a natural number object for any small $\mathcal{C}$, namely the constant functor with value $\mathbb{N}$: the functor $\Delta: \mathbf{Set} \to [\mathcal{C}, \mathbf{Set}]$ sending a set to the corresponding constant functor preserves 1 and has a right adjoint, which sends a functor $\mathcal{C} \to \mathbf{Set}$ to its limit. Again, if N is a natural number object in a properly cartesian closed category $\mathcal{C}$ then $A^*(N)$ is a natural number object in $\mathcal{C}/A$ for any A, by 1.5.2(i) (cf. 2.5.3 above).

In any category with a terminal object and countable coproducts, it is easy to see that the countable copower $\sum_{n \in \mathbb{N}} 1$ satisfies the condition of 2.5.1. This provides an alternative proof of 2.5.6(i) for functors between categories with countable coproducts. However, even in a topos a natural number object need not be a countable copower of 1; see, for example, D5.1.7 below, or alternatively F2.1.6(iv).

For future reference, we note the following result about countable colimits:

Lemma 2.5.7 *Let $\mathcal{C}$ be a σ-pretopos (as defined at the end of Section* A1.4*), $\mathcal{E}$ a topos with countable coproducts and $F: \mathcal{C} \to \mathcal{E}$ a coherent functor. Then F is σ-coherent (i.e. preserves countable coproducts) iff it is cocartesian (i.e. preserves coequalizers).*

Proof One direction follows directly from 1.4.19: if F is σ-coherent, then it preserves all the structure involved in the construction of coequalizers in $\mathcal{C}$. Conversely, suppose F preserves coequalizers. Then it preserves the natural number object, by 2.5.6(ii); equivalently, by the discussion above, it preserves the countable copower of 1. But, given any countable family $(A_n \mid n \in \mathbb{N})$ of objects of $\mathcal{C}$, the coproduct $A = \coprod_{n \in \mathbb{N}} A_n$ is characterized up to isomorphism by the existence of morphisms $\nu_n: A_n \to A$ and $\rho: A \to N$ making

$$\begin{tikzcd} A_n \arrow[r, "\nu_n"] \arrow[d] & A \arrow[d, "\rho"] \\ 1 \arrow[r, "\nu_n"] & N \end{tikzcd}$$

a pullback for each n; F preserves these pullbacks, and hence preserves the coproduct of the A_n. □

Next, we note a useful consequence of Lemma 2.5.5(i):

Corollary 2.5.8 *Let (N, o, s) be a natural number object in a positive coherent category. Then the monomorphisms $o: 1 \rightarrowtail N$ and $s: N \rightarrowtail N$ are complementary in* $\mathrm{Sub}(N)$. □

In particular, we note that the subobjects o and s are disjoint – that is, Peano's third postulate is satisfied. To complete the discussion of the Peano postulates, we have

Lemma 2.5.9 *In any category, a natural number object satisfies Peano's fifth postulate: i.e. if $N' \rightarrowtail N$ is a subobject such that both $o: 1 \to N$ and the restriction of s to N' factor through N', then N' is the whole of N.*

Proof Let $o': 1 \to N'$ and $s': N' \to N'$ be the factorizations. Then there exists $f: N \to N'$ such that

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "o'"'] & N \arrow[r, "s"] \arrow[d, "f"] & N \arrow[d, "f"] \\ & N' \arrow[r, "s'"] & N' \end{tikzcd}$$

commutes; and the composite $N \to N' \rightarrowtail N$ must be the identity since it commutes with both o and s. So $N' \rightarrowtail N$ is split epic, and hence an isomorphism. □

Once again, we shall see in Section D5.1 that the Peano postulates suffice to characterize the natural number object in a topos.

Until further notice, we shall work in a cartesian closed, positive Heyting category $\mathcal{E}$. One important consequence of 2.5.8 is that, since complements are preserved under pullback, it allows us to construct lots of other complemented subobjects of (finite powers of) N. For example, we define the order-relation $N_1 \rightarrowtail N \times N$ by the pullback

$$\begin{tikzcd} N_1 \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "o"] \\ N \times N \arrow[r, "\dot{-}"] & N \end{tikzcd}$$

Of course, it requires proof that this is an order relation on N; we shall indicate briefly how this may be done. First we note that, since the identity $n \dot{-} n = 0$ is

provable in primitive recursive arithmetic, it is valid in our context; that is, the diagonal $\Delta \colon N \rightarrowtail N \times N$ factors through N_1.

Proposition 2.5.10 *The intersection of $N_1 \rightarrowtail N \times N$ and the opposite relation $N_1^{\circ} \rightarrowtail N \times N$ is the diagonal.*

Proof After the preceding remark, we have to show that $N_1 \cap N_1^{\circ} \leq \Delta$, or equivalently that $((N_1 \cap N_1^{\circ}) \Rightarrow \Delta)$ is the whole of $N \times N$. Let $N' \rightarrowtail N$ be the image of the latter subobject under $\forall_{\pi_1} \colon \mathrm{Sub}(N \times N) \to \mathrm{Sub}(N)$ (in set-theoretic terms, N' is the set

$$\{m \in N \mid (\forall n \in N)((m \dot{-} n = 0 = n \dot{-} m) \Rightarrow (m = n))\});$$

it suffices to show that N' is the whole of N, and we shall do this using 2.5.9. Firstly, o factors through N', since $n \dot{-} 0 = n$, and so the intersection of $o \times 1 \colon 1 \times N \rightarrowtail N \times N$ with N_1° is just $(o, o) \colon 1 \rightarrowtail N \times N$. To show that s restricts to a morphism $N' \to N'$, we shall argue informally in terms of elements (the logical underpinnings of this informal argument will be found in Section D1.3). Suppose $m \in N'$, and suppose $sm \dot{-} n = 0 = n \dot{-} sm$. By the first equation, $n \neq 0$, so n is a successor, say $n = sk$. But $sm \dot{-} sk = m \dot{-} k$ for any m and k, so we deduce $m \dot{-} k = 0 = k \dot{-} m$, whence $m = k$ (since $m \in N'$) and so $sm = sk = n$. Thus $sm \in N'$. □

Corollary 2.5.11 *In any topos, the natural number object (if it exists) is decidable (cf. 1.4.15); i.e. the diagonal subobject $N \rightarrowtail N \times N$ is complemented in* $\mathrm{Sub}(N \times N)$.

Proof N_1 is complemented in $N \times N$ because it is a pullback of $o \colon 1 \rightarrowtail N$, which is complemented by 2.5.8. But a finite intersection of complemented subobjects is complemented, in any coherent category. □

Proposition 2.5.12 *The union of N_1 and N_1° is the whole of $N \times N$.*

Proof Let $E \rightarrowtail N \times N$ be the equalizer of $\pi_1 \colon N \times N \to N$ and

$$N \times N \xrightarrow{(\dot{-}, \pi_2)} N \times N \xrightarrow{+} N$$

(i.e. the set $\{(m, n) \mid (m \dot{-} n) + n = m\}$); we shall show that $N_1 \cup E$ is the whole of $N \times N$ and that $E \leq N_1^{\circ}$. (Note, incidentally, that E is a complemented subobject of $N \times N$, since it is expressible as a pullback of the diagonal; but we shall not need this fact.)

For the first assertion, let $N' \rightarrowtail N = \forall_{\pi_2}(N_1 \cup E \rightarrowtail N \times N)$; we shall argue as in the proof of 2.5.10 to show that N' is the whole of N. It contains o (in fact $\forall_{\pi_2}(E)$ contains o, since $m \dot{-} 0 = m + 0 = m$). If $n \in N'$, then for any $m \in N$ we have either $m \dot{-} n = 0$, whence $m \dot{-} sn = p0 = 0$, or $(m \dot{-} n) + n = m$. But in

the latter case we have

$$(m \dot{-} sn) + sn = s(p(m \dot{-} n) + n) = sp(m \dot{-} n) + n,$$

and $m \dot{-} n$ is either 0 (which reduces to the previous case) or a successor (in which case $sp(m \dot{-} n) = m \dot{-} n$, so $(m \dot{-} sn) + sn = m$). Thus $sn \in N'$.

For the second assertion, it suffices to observe that $n \dot{-} (n + k) = 0$ for all n and k, so if $(m, n) \in E$ we have

$$n \dot{-} m = n \dot{-} (n + (m \dot{-} n)) = 0,$$

i.e. $(m, n) \in N_1^{\circ}$. □

Proposition 2.5.13 *The relation $N_1 \rightarrowtail N \times N$ is transitive (in the sense of 1.3.6(d)).*

Proof Once again, we argue by induction, using 2.5.9. Let $N' \rightarrowtail N$ be

$$\forall_{\pi_1}((\pi_{12}^*(N_1) \cap \pi_{23}^*(N_1)) \Rightarrow \pi_{13}^*(N_1)),$$

i.e. the set $\{m \in N \mid (\forall n, k \in N)((m \dot{-} n = n \dot{-} k = 0) \Rightarrow (m \dot{-} k = 0))\}$. We have $o \in N'$, since $0 \dot{-} k = 0$ for all k. Suppose $m \in N'$, and suppose $sm \dot{-} n = n \dot{-} k = 0$. Then $n \neq 0$, so n must be a successor (say $n = sn'$), and similarly $k = sk'$. Now we have $m \dot{-} n' = sm \dot{-} sn' = 0$ and similarly $n' \dot{-} k' = 0$, so $m \dot{-} k' = 0$; but this implies $sm \dot{-} k = 0$. So $sm \in N'$. □

Putting together the last four results, we have shown that $N_1 \rightarrowtail N \times N$ is a (trichotomous) linear order on N, in any cartesian closed positive Heyting category $\mathcal{E}$ with a natural number object.

Next, we consider the composite

$$N_1 \rightarrowtail N \times N \xrightarrow{\pi_2} N \xrightarrow{s} N$$

as an object of $\mathcal{E}/N$. We denote this object by C, and call it the *generic finite cardinal* in $\mathcal{E}$. If $p: A \to N$ is any morphism with codomain N (which we think of as a 'natural number in the category $\mathcal{E}/A$' – recall that $A^*(N)$ is a natural number object in this category, by 2.5.6(i)), we shall write $[p]$ for the pullback of C along p, and call the objects which arise in this way *finite cardinals* in $\mathcal{E}/A$. (Thus C itself could alternatively be denoted $[1_N]$.) The reason for this terminology is that, in the case $\mathcal{E} = \mathbf{Set}$, C is the $\mathbb{N}$-indexed family whose pth member is the set $\{0, 1, 2, \ldots, p-1\}$, i.e. it is the cardinal p itself as defined by von Neumann. In general, we have

Lemma 2.5.14 *$[o] \cong 0$ and $[sp] \cong 1 \amalg [p]$ for any p.*

Proof The first assertion follows from the disjointness of o and s (2.5.8). For the second, it suffices to prove the result for $p = 1_N$, since pullback functors preserve coproducts. But since s is monic, $[s]$ is the object $\pi_2: N_1 \rightarrowtail N \times N \to N$.

And we know that, as a subobject of $N \times N$, N_1 is the coproduct of the diagonal $N \rightarrowtail N \times N$ and the strict order relation (the complement of N_1°); and the latter may easily be identified with

$$N_1 \rightarrowtail N \times N \overset{1 \times s}{\rightarrowtail} N \times N$$

by an argument like those already given. Composing with π_2, we obtain a coproduct decomposition $[s] \cong 1_N \amalg C$ in $\mathcal{E}/N$. □

Further properties of finite cardinals will be studied in Section D5.2.

We conclude this section with a brief discussion of list objects.

Definition 2.5.15 A *list object* over an object A in a category $\mathcal{E}$ with finite products is an object LA equipped with morphisms $o_A : 1 \to LA$ and $s_A : A \times LA \to LA$ which is initial in the category of objects equipped with such structure; i.e. given any object B with morphisms $b: 1 \to B$ and $t: A \times B \to B$, there is a unique $f : LA \to B$ such that

$$\begin{tikzcd} 1 \arrow[r, "o_A"] \arrow[dr, "b"'] & LA \arrow[d, "f"] & A \times LA \arrow[l, "s_A"'] \arrow[d, "1 \times f"] \\ & B & A \times B \arrow[l, "t"] \end{tikzcd}$$

commutes.

The notation $A*$ is also commonly used for the list object over A. Clearly, in the case $A = 1$ the definition reduces to that of a natural number object; we can think of list objects as standing in the same relation to natural number objects as power objects do to subobject classifiers. In **Set**, we may take LA to be the set of all finite lists of elements of A, with o_A picking out the empty list and s_A corresponding to the operation of appending an element of A to the head of a list. We note that, if $\mathcal{E}$ has list objects, then the operation $A \mapsto LA$ becomes a functor $\mathcal{E} \to \mathcal{E}$: given $f : A \to B$, we define Lf to be the unique morphism making

$$\begin{tikzcd} 1 \arrow[r, "o_A"] \arrow[dr, "o_B"'] & LA \arrow[d, "Lf"] & A \times LA \arrow[l, "s_A"'] \arrow[d, "1 \times Lf"] \\ & LB & A \times LB \arrow[l, "s_B(f \times 1)"] \end{tikzcd}$$

commute. In particular, any list object comes equipped with a morphism $l: LA \to L1 = N$ which 'measures the length of a list'.

Many of the formal properties of natural number objects, established in this section, are inherited by list objects. For example, in analogy with 2.5.5, we have a coproduct decomposition $LA \cong 1 + (A \times LA)$ provided coproducts exist in $\mathcal{E}$, and the coequalizer of s_A and $\pi_2 \colon A \times LA \to LA$ is $LA \to 1$. Similarly, generalizing 2.5.6(i), if $F \colon \mathcal{E} \to \mathcal{F}$ is a functor which preserves finite products and has a right adjoint, and LA is a list object over $A \in \mathrm{ob}\,\mathcal{E}$, then $F(LA)$ is a list object over FA. We omit the proofs, which are straightforward generalizations of those already given.

In establishing the existence of list objects in suitable categories, we shall need to use the following 'section criterion' due to G. C. Wraith [549].

Lemma 2.5.16 *Let $\mathcal{E}$ be a cartesian category having a natural number object N, and let $f \colon A \to N$ be an object of $\mathcal{E}/N$. If we are given morphisms $a \colon 1 \to o^*(f)$ in $\mathcal{E}$ and $t \colon f \to s^*(f)$ in $\mathcal{E}/N$, then there exists a unique $h \colon 1_N \to f$ in $\mathcal{E}/N$ such that $o^*(h) = a$ and $s^*(h) = th$.*

Proof Consider the diagram

$$\begin{tikzcd}
1 \arrow[r, "o"] \arrow[d, "a"] & N \arrow[rr, "s"] \arrow[d, "h"] & & N \arrow[d, "h"] \\
o^*(A) \arrow[r] \arrow[d] & A \arrow[r, "t"] \arrow[d, "f"'] & s^*(A) \arrow[r] \arrow[dl] & A \arrow[d, "f"] \\
1 \arrow[r, "o"] & N \arrow[rr, "s"] & & N
\end{tikzcd}$$

in which the morphism h exists uniquely by the definition of N, and the composite fh is the identity by the uniqueness clause of 2.5.1. Moreover, a standard diagram-chase shows that the upper left square of the diagram and the square

$$\begin{tikzcd}
N \arrow[r, "s"] \arrow[d, "th"] & N \arrow[d, "t"] \\
s^*(A) \arrow[r] & A
\end{tikzcd}$$

are pullbacks. □

Theorem 2.5.17 *Let $\mathcal{E}$ be a locally cartesian closed, positive coherent category with a natural number object. Then $\mathcal{E}$ has list objects.*

Proof We define LA to be the partial product $P(A, C)$ as defined in 1.5.7, i.e. the object $\Sigma_N((N^*A)^C)$, where the exponential is calculated in $\mathcal{E}/N$; and we shall write $l\colon LA \to N$ for the object $(N^*A)^C$ itself. We note first that, by 2.5.14 and the fact that pullback functors preserve exponentials (1.5.2(ii)), we have $o^*(l) \cong A^{[o]} \cong A^0 \cong 1$ and $s^*(l) \cong (N^*A)^{[s]} \cong N^*A^{(1_N \amalg C)} \cong N^*A \times l$; so the pullbacks of o and s along l provide the required structure maps $1 \to LA$ and $A \times LA \to LA$.

Now suppose given an object B with morphisms $b\colon 1 \to B$ and $t\colon A \times B \to B$. We need to construct a morphism $\Sigma_N(N^*A^C) \to B$ in $\mathcal{E}$, or equivalently a morphism $1_N \to (N^*B)^{(N^*A^C)}$ in $\mathcal{E}/N$. But by 2.5.14 we can construct this from a morphism $1 \to o^*((N^*B)^{(N^*A^C)})$ in $\mathcal{E}$ and a morphism $(N^*B)^{(N^*A^C)} \to s^*((N^*B)^{(N^*A^C)})$ in $\mathcal{E}/N$. And, again using the fact that pullback functors preserve exponentials, we have $o^*((N^*B)^{(N^*A^C)}) \cong B^{(A^0)} \cong B^1 \cong B$, so we may take the former to be the given morphism b; and $s^*((N^*B)^{(N^*A^C)}) \cong (N^*B)^{(N^*A \times N^*A^C)}$, so for the latter we take the transpose (u, say) of

$$(N^*B)^{(N^*A^C)} \times N^*A \times N^*A^C \xrightarrow{(\pi_2, \mathrm{ev} \circ \pi_{13})} N^*A \times N^*B \xrightarrow{N^*t} N^*B .$$

Moreover, a straightforward diagram-chase shows that a morphism $LA \to B$ satisfies the conditions of 2.5.15 iff its transpose $h\colon 1_N \to (N^*B)^{(N^*A^C)}$ satisfies $o^*(h) = b$ and $s^*(h) = uh$, so this is the unique morphism with the required property. □

For a simpler, but less explicit, proof of 2.5.17 in the case when $\mathcal{E}$ is a topos, see D5.1.5. Further properties of natural number objects and of list objects will be explored in Chapter D5.

Suggestions for further reading: Freyd [371], Johnstone & Wraith [549].

A2.6 Quasitoposes

Although one great merit of the topos axioms is their wide applicability, as demonstrated by the list of examples in Section A2.1, there are certain contexts in which they appear inconveniently restrictive. The problem arises from Corollary 1.6.2, which told us that a category with a subobject classifier (in particular, a topos) is necessarily balanced, whereas we do sometimes encounter examples of non-balanced categories which are sufficiently 'topos-like' to share many important properties with them. The notion of quasitopos, introduced by J. Penon [954,956], succeeds in capturing these examples (as well as all toposes); we devote this section to studying it.

Clearly, one way of weakening the notion of subobject classifier is to demand classifying maps not for arbitrary subobjects but only for the members of some restricted class of subobjects (which might as well be stable under pullback).

Whatever class we choose, it will follow from the argument of 1.6.2 that all monomorphisms in this class are regular; so one possibility would be to take the class of regular monomorphisms. However, for axiomatizing the notion of quasitopos it is more convenient to take the class of cocovers (more usually called strong monomorphisms); we shall then obtain for free the information that every cocover is regular monic (compare 1.3.4, where a substantial argument was required to establish the dual result in a regular category).

Accordingly, we define a *weak subobject classifier* in a cartesian category $\mathcal{E}$ to be an object Ω equipped with a morphism $\top : 1 \to \Omega$ such that, given any cocover $A' \rightarrowtail A$ in $\mathcal{E}$, there is a unique $f : A \to \Omega$ making

$$\begin{tikzcd} A' \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\ A \arrow[r, "f"] & \Omega \end{tikzcd}$$

a pullback. (There is no danger in re-using the notation of Section A1.6, since a subobject classifier – if it exists – is necessarily a weak subobject classifier.)

Definition 2.6.1 A *quasitopos* is a locally cartesian closed, cocartesian category with a weak subobject classifier.

We note that, in compensation for weakening the notion of subobject classifier, we have had to strengthen our axioms in two other respects. First, we have had to assume the existence of finite colimits, instead of proving it as in Section A2.2: we may, of course, define the *weak power object* PA of an object A to be Ω^A, and as previously we can make P into a functor $\mathcal{E}^{\mathrm{op}} \to \mathcal{E}$; but this functor is no longer monadic, because (although it is still faithful, by the argument of 2.2.3(ii)) it fails to be conservative. (In fact we shall see that if f is any morphism which is both monic and epic in a quasitopos, then Pf is an isomorphism.) Secondly, we have had to strengthen 'properly cartesian closed' to 'locally cartesian closed', in order to ensure that the property of being a quasitopos is stable under slicing: although we can argue just as in 2.3.2 to show that the possession of weak power objects is inherited by $\mathcal{E}/B$ from $\mathcal{E}$, we cannot reconstruct exponentials from weak power objects. (The point at which the construction of 2.4.11 breaks down is that, although we can define the singleton map $\{\} : A \to PA$ (since the diagonal $A \rightarrowtail A \times A$ is regular monic, being the equalizer of the two projections) and we can prove as in 2.2.3(i) that $\{\}$ is monic, it is not a cocover in general; hence (in the notation of 2.4.11) the relation $Q \rightarrowtail P(A \times B) \times A$ need not be a cocover, and so need not have a name.)

Despite the foregoing, it follows immediately from 2.2.9 and 2.3.4 that a topos is a quasitopos. Conversely, we shall see shortly that a balanced quasitopos is a topos. First we need a quasitopos version of 2.4.3:

Lemma 2.6.2 *Given a pushout square*

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, tail, "f"'] & C \arrow[d, "d"] \\ B \arrow[r, "e"] & D \end{tikzcd}$$

in a quasitopos where f is a cocover, the morphism d is also a cocover and the square is also a pullback.

Proof The morphism $(f, g)\colon A \rightarrowtail B \times C$ is a cocover, since f is (this is where we use the fact that we defined weak subobject classifiers in terms of cocovers rather than regular monomorphisms), and so is named by a morphism $h\colon B \to PC$. Exactly as in 2.4.2, we may show that the square

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, tail, "f"'] & C \arrow[d, tail, "\{\}"] \\ B \arrow[r, "h"] & PC \end{tikzcd}$$

(commutes and) is a pullback; hence, as in 2.4.3, the pair $(h, \{\})$ factors through (e, d), and the square in the statement of the lemma is a pullback. We may also conclude that d is monic, since $\{\}$ is; but, as indicated above, we must do some further work to prove that d is a cocover.

Let $k\colon B \to \Omega$ be the classifying map of f; then the diagram

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, tail, "f"'] & C \arrow[r] & 1 \arrow[d, tail, "\top"] \\ B \arrow[rr, "k"] & & \Omega \end{tikzcd}$$

commutes (since both ways round classify the top element of $\mathrm{Sub}(A)$), and hence induces a morphism $l\colon D \to \Omega$. Let $d'\colon C' \rightarrowtail D$ be the (regular) subobject classified by l; then $d' \geq d$ in $\mathrm{Sub}(D)$, and the pullback of d' along e is isomorphic to f. We shall show that the factorization $C \rightarrowtail C'$ of d through d' is an isomorphism,

so that d is a cocover. Since D is a pushout, we have a regular epimorphism $B + C \to D$; pulling this back along d', we get a regular epimorphism (in particular, a cover) $A + C \twoheadrightarrow C'$, since pullback functors preserve coproducts and regular epimorphisms. But $A \to C'$ factors through $C \to C'$; so the latter must be a cover, and hence an isomorphism since we already know it is monic. □

Corollary 2.6.3

(i) *A quasitopos is coregular.*

(ii) *A balanced quasitopos is a topos.*

Proof (i) This is immediate from 2.6.2 and 1.3.5, since we already know that cocovers coincide with regular monomorphisms.

(ii) By (i), every monomorphism in a quasitopos factors as a monic epimorphism followed by a cocover; if the quasitopos is balanced, the first factor is an isomorphism, and so every monomorphism is a cocover. Thus the weak subobject classifier is a subobject classifier. □

Before we develop the theory of quasitoposes any further, we should look at some examples which are not toposes.

Examples 2.6.4 (*a*) Any Heyting algebra (more generally, any cartesian closed preorder which is cocartesian) is a quasitopos. We saw in Section A1.5 that such a category is locally cartesian closed; and since its only cocovers are isomorphisms, the terminal object will serve as a weak subobject classifier.

(*b*) We have already noted that **Sp** is not cartesian closed; however, various constructions have been proposed for embedding **Sp** in a cartesian closed category $\mathcal{C}$. It is clear that if such an embedding is reasonably well-behaved (at least if it preserves monomorphisms and epimorphisms and reflects isomorphisms) then $\mathcal{C}$ cannot be a topos, since it will inherit non-balancedness from **Sp**; but we can hope for it to be a quasitopos, and in fact several of the proposals put forward for $\mathcal{C}$ are quasitoposes. One such, which is canonical in a certain sense (see [1247]), is the category **Choq** of Choquet spaces (convergence spaces in the sense of Choquet [240]), and we now describe this example. A *Choquet space* is a set X equipped with a binary relation between (proper) filters Φ on X and points $x \in X$, written $\Phi \to x$ (and read 'Φ converges to x'), such that

(i) For every $x \in X$, the principal ultrafilter $\{X' \subseteq X \mid x \in X'\}$ converges to x.

(ii) If $\Phi \to x$ and $\Psi \supseteq \Phi$, then $\Psi \to x$.

(iii) If Φ is a filter and x a point such that every filter containing Φ is itself contained in a filter converging to x, then $\Phi \to x$.

A *continuous map* of Choquet spaces is a function between their underlying sets which preserves convergence, in the obvious sense; this defines the category **Choq**. The usual (Cartan [235]) notion of convergence for filters on a topological space defines a functor **Sp** $\to$ **Choq**, which is well known to be full and faithful.

It is straightforward to verify that **Choq** has finite limits and colimits, and that the forgetful functor to **Set** preserves them; in fact the latter has both left and right adjoints, which respectively equip a set X with the *discrete* Choquet structure in which only principal ultrafilters converge, and with the *indiscrete* structure in which every filter converges to every point. The monomorphisms (resp. epimorphisms) in **Choq** are just the injective (resp. surjective) continuous maps, from which it follows that the cocovers are just those injections $i: Y \to X$ for which Y has the 'subspace Choquet structure' (i.e. every filter on Y which maps to a filter converging in X to $i(y)$ converges in Y to y), and hence that the indiscrete two-point Choquet space is a weak subobject classifier. We shall not give the proof that **Choq** is locally cartesian closed, since it would occupy too much space – for more details of this and related examples, see the papers of Wyler [1247] and Dubuc [297].

(*c*) There is a sequential version of the last example, which essentially goes back to M. Fréchet [369]. By a *Fréchet space* we mean a set X equipped with a relation of convergence between sequences in X and points of X, such that

(i) For every $x \in X$, the constant sequence (x) converges to x.
(ii) If $(x_n) \to x$, then every subsequence of (x_n) converges to x.
(iii) If (x_n) is a sequence and x a point such that every subsequence of (x_n) contains a subsequence converging to x, then $(x_n) \to x$.

(Fréchet's original definition also included the 'separation axiom' that every sequence has at most one limit, but for our present purposes it is essential to omit this requirement, since we need to have the indiscrete spaces in our category.) The category of Fréchet spaces (with convergence-preserving functions as morphisms) will be denoted **Fre**; the proof that it is a quasitopos is similar to that for **Choq**, and we omit the details. The usual notion of sequential convergence in topological spaces gives a functor $\mathbf{Sp} \to \mathbf{Fre}$, which is faithful but not conservative; however, it is full when restricted to the subcategory of sequential spaces (cf. Example 2.1.11(*j*)). Moreover, since the space $\mathbb{N}^+$ is sequential, it is easy to verify that the functor $\mathbf{Sp}(\mathbb{N}^+, -): \mathbf{Sp} \to \mathbf{Sh}(M)$ of 2.1.11(*j*) extends to a functor $\mathbf{Fre}(\mathbb{N}^+, -): \mathbf{Fre} \to \mathbf{Sh}(M)$; the latter is full and faithful (note that continuous maps $\mathbb{N}^+ \to X$ in **Fre** correspond to instances of the convergence relation in X). For more discussion of this example, see [507], and also C2.2.14(*b*).

(*d*) Let $(\mathcal{C}, T)$ be a site, as defined in 2.1.9. We say a functor $F: \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is *separated* if it satisfies the 'uniqueness' but not the 'existence' part of the sheaf axiom for T-covering families; equivalently, if for every $(f_i: A_i \to A \mid i \in I) \in T(A)$ the function

$$F(A) \longrightarrow \prod_{i \in I} F(A_i)$$

induced by the $F(f_i)$ is injective. We write $\mathbf{Sep}(\mathcal{C}, T) \subseteq [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ for the full subcategory of T-separated functors; we claim that $\mathbf{Sep}(\mathcal{C}, T)$ is a quasitopos.

The proof that $\mathbf{Sep}(\mathcal{C}, T)$ is closed under finite limits and an exponential ideal in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is similar to that for $\mathbf{Sh}(\mathcal{C}, T)$. It is not closed under finite colimits, but it has them by virtue of the fact that it is reflective in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; we shall not prove this here, since it is a special case of a general result which we shall establish in Section A4.4. For local cartesian closedness, we recall from 1.1.7 that any slice $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]/F$ is equivalent to $[\mathcal{F}^{\mathrm{op}}, \mathbf{Set}]$ for a suitable category $\mathcal{F}$; and we can 'lift' the coverage T on $\mathcal{C}$ to a coverage T_F on $\mathcal{F}$, in such a way that $\mathbf{Sep}(\mathcal{C}, T)/F$ is identified with $\mathbf{Sep}(\mathcal{F}, T_F)$. Finally, a monomorphism $F' \rightarrowtail F$ in $\mathbf{Sep}(\mathcal{C}, T)$ is easily seen to be a cocover iff it has the property that, given any $A \in \mathrm{ob}\,\mathcal{C}$ and any $s \in F(A)$, if there exists a covering family $(f_i : A_i \to A \mid i \in I)$ such that $F(f_i)(s) \in F'(A_i)$ for each $i \in I$, then $s \in F'(A)$. Using this, it is easy to see that the subobject classifier Ω_T of $\mathbf{Sh}(\mathcal{C}, T)$, as constructed in 2.1.10, is a weak subobject classifier for $\mathbf{Sep}(\mathcal{C}, T)$.

(*e*) Our next example is a modification of the category $\mathbf{Pos}(L)$, for a distributive lattice L, which we described in 1.4.6. Suppose that L is *irreducible*, i.e. that the meet of any two nonzero elements of L is nonzero. Then since any sequence of elements of L, regarded as an object of $\mathbf{Pos}(L)$, is isomorphic to the sequence obtained by deleting all occurrences of 0 from the original one, and since the nonzero elements of L have no nontrivial disjoint decompositions, we may simplify the description of $\mathbf{Pos}(L)$ as we did for the particular case $L = \mathbf{2}$ in 1.4.6 itself: it is equivalent to the concrete category whose objects are finite sets A equipped with a function $\alpha : A \to L \setminus \{0\}$, and whose morphisms $f : (A, \alpha) \to (B, \beta)$ are functions satisfying $\alpha(a) \leq \beta(f(a))$ for all $a \in A$. We denote this category by $\mathbf{Fuz}_f(L)$, and we claim that *if L is an irreducible Heyting algebra then $\mathbf{Fuz}_f(L)$ is a quasitopos*. It is easy to verify that $\mathbf{Fuz}_f(L)$ has finite limits and colimits (and the forgetful functor $\mathbf{Fuz}_f(L) \to \mathbf{Set}_f$ preserves them). It is cartesian closed: $(B, \beta)^{(A,\alpha)}$ may be taken to be the set F of all functions $f : A \to B$ with structure map $\phi : F \to L \setminus \{0\}$ given by

$$\phi(f) = \bigwedge \{\alpha(a) \Rightarrow \beta f(a) \mid a \in A\}$$

(note that this meet is never 0, since A is finite and L is irreducible), and the exponential adjunction is easy to verify. Exponentials in slices of $\mathbf{Fuz}_f(L)$ are similarly defined. Cocovers in $\mathbf{Fuz}_f(L)$ are injective functions for which the inequality in the definition of a morphism becomes an equality; from this one sees easily that a two-element set, with α sending both its elements to 1, is a weak subobject classifier.

In the particular case $L = \mathbf{2}$, this construction in fact yields a topos, as we saw in 1.4.6; but in general it does not – for example, if L is a totally ordered set with more than two elements, then we saw after 1.4.6 that $\mathbf{Pos}(L)$ is not effective as a regular category, and so it cannot be a topos. If L is a complete irreducible Heyting algebra, then we may embed $\mathbf{Fuz}_f(L)$ in a larger quasitopos $\mathbf{Fuz}(L)$, simply by removing the finiteness restriction on the underlying sets of its objects (and by modifying the definition of exponentials: we have to cut down

to the subset $F' \subseteq F$ of functions f such that $\phi(f) \neq 0$). This category is one version of the category of *L-valued fuzzy sets*, as studied by fuzzy-set theorists; another version in common use, which allows the possibility $\alpha(a) = 0$, may be identified with $\mathbf{Fuz}(L_\perp)$, where $L_\perp$ is the lattice obtained from L by adjoining a new bottom element below the original one (note that such a lattice is always irreducible). Of course, we may define the categories $\mathbf{Fuz}_f(L)$ and $\mathbf{Fuz}(L)$ for any lattice L, irreducible or not; but for Heyting algebras which are not irreducible we do not get a cartesian closed category. (The reason is that the definition of products has to be modified: in forming the product of (A, α) and (B, β) we have to exclude the pairs (a, b) with $\alpha(a) \wedge \beta(b) = 0$. See also [1248, Chapter 8], for a more detailed discussion of the categories which have been proposed by fuzzy-set theorists.)

(f) One further example, which we shall not discuss here, but which will be of importance in Chapter F2, is the category $\mathbf{Ass}(A)$ of A-valued assemblies, where A is a partial Schönfinkel algebra – see F2.1.6.

We now revert to the general theory of quasitoposes. As well as being coregular, a quasitopos is regular (in fact, a Heyting category) by 1.5.13; it need not be effective or positive (indeed, if it has both properties then it is a topos, by 1.4.9 and 2.6.3(ii)), but it can have either property without having the other. For example, a Heyting algebra is (trivially) effective, but not positive unless it is degenerate, while the quasitoposes **Choq** and **Fre** are positive but not effective. Regarding positivity, we have the following result:

Lemma 2.6.5 *A quasitopos has disjoint coproducts iff the unique morphism* $0 \to 1$ *is a cocover.*

Proof If $0 \to 1$ is a cocover, then 0 cannot have any proper epimorphic images, and disjointness of coproducts follows from 1.5.14 (or alternatively it can be deduced from 2.6.2 just as we deduced 2.4.4 from 2.4.3). Conversely, if $1 + 1$ is disjoint, then $0 \to 1$ is an equalizer of the two coprojections $1 \rightrightarrows 1 + 1$. □

Quasitoposes satisfying the conditions of 2.6.5 have been called *solid* by G. P. Monro [861]. Let $\mathcal{E}$ be a general quasitopos, and let $0 \twoheadrightarrow \overline{0} \rightarrowtail 1$ be the factorization of $0 \rightarrowtail 1$ into a (monic) epimorphism followed by a cocover. It is easy to see that the analogue of 1.6.7 holds for weak subobject classifiers, and hence that any slice category of a quasitopos is a quasitopos; in particular, $\mathcal{E}/\overline{0}$ is a quasitopos. But it is also a preorder; in general, any cartesian closed category in which $0 \to 1$ is epic must be a preorder, since two distinct morphisms $A \rightrightarrows B$ would yield distinct morphisms $1 \rightrightarrows B^A$. Moreover, since $\overline{0}$ is a subobject of 1, we may identify $\mathcal{E}/\overline{0}$ with a full subcategory of $\mathcal{E}$, whose objects are those objects of $\mathcal{E}$ which admit a morphism to $\overline{0}$ (equivalently, those subobjects of 1 contained in $\overline{0}$); we call these *thin* objects of $\mathcal{E}$. More surprisingly, we have

Lemma 2.6.6 *With the above notation, the co-slice category* $\overline{0}\backslash\mathcal{E}$ *is a positive quasitopos.*

Proof Again, since $0 \to \overline{0}$ is epic, we may identify $\overline{0}\backslash\mathcal{E}$ with a full subcategory of $\mathcal{E}$, consisting of those objects which admit a morphism from $\overline{0}$; we call these *solid* objects. As such, it is easily seen to be closed under finite limits and pushouts, and to have an initial object $1_{\overline{0}}$; and it is an exponential ideal since, for any two objects A and B, there exists a morphism $B \to B^A$, namely the transpose of the first projection. It is locally cartesian closed, because slicing commutes with co-slicing: for any solid object A, we have

$$(\overline{0}\backslash\mathcal{E})/(\overline{0} \to A) \cong (\overline{0} \to A)\backslash(\mathcal{E}/A),$$

but $(\overline{0} \to A)$ is the '$\overline{0}$-object' of $\mathcal{E}/A$, and so the category on the right is cartesian closed. And $\overline{0}\backslash\mathcal{E}$ inherits a weak subobject classifier from $\mathcal{E}$, by the argument of 1.6.8. Finally, $\overline{0} \to 1$ is a cocover in $\overline{0}\backslash\mathcal{E}$, since it is one in $\mathcal{E}$; so by 2.6.5 $\overline{0}\backslash\mathcal{E}$ is positive. □

Thus, in a sense, the study of quasitoposes 'bifurcates' into the study of those which are preorders and those which are positive. By this we do not mean to imply that, given an arbitrary quasitopos $\mathcal{E}$, we can reconstruct it up to equivalence from the preorder $\mathcal{P} = \mathcal{E}/\overline{0}$ and the positive quasitopos $\mathcal{Q} = \overline{0}\backslash\mathcal{E}$: there are at least two ways of putting these together, namely the product $\mathcal{P} \times \mathcal{Q}$ and the ordinal sum $\mathcal{P} \oplus \mathcal{Q}$ (the latter being obtained from the disjoint union by identifying the terminal object of $\mathcal{P}$ with the initial object of $\mathcal{Q}$), and these are distinct provided both $\mathcal{P}$ and $\mathcal{Q}$ are non-degenerate. However, it turns out that one further piece of information suffices to recover $\mathcal{E}$ from $\mathcal{P}$ and $\mathcal{Q}$:

Proposition 2.6.7 *Let $\mathcal{E}$ be a quasitopos, and let $\mathcal{P}$ and $\mathcal{Q}$ be defined as above. Then there is a cartesian functor $F\colon \mathcal{P} \to \mathcal{Q}$ such that $\mathcal{E}$ is equivalent to the category* $\mathbf{Gl}(F)$ *obtained by glueing along F, as in* 2.1.12.

In passing, we remark that the glueing construction applied to a cartesian functor between quasitoposes always yields a quasitopos; this can be proved directly by methods like those of 2.1.12, but will follow from a more general result to be proved in Section A4.2.

Proof To simplify the notation, we identify $\mathcal{P}$ and $\mathcal{Q}$ with the appropriate full subcategories of $\mathcal{E}$. For a thin object A, we define $F(A) = A^{\overline{0}} + \overline{0}$; we note that since $A^{\overline{0}}$ is subterminal by 1.5.10, we have

$$A^{\overline{0}} \times \overline{0} \cong A \times \overline{0} \cong A$$

by one of the Heyting semilattice identities of 1.5.11 and the fact that $A \leq \overline{0}$ in $\mathrm{Sub}(1)$. And since $0 \to A$ is epic (being a pullback of $0 \to \overline{0}$), the square

$$\begin{array}{ccc} A & \longrightarrow & \overline{0} \\ \downarrow & & \downarrow \\ A^{\overline{0}} & \longrightarrow & F(A) \end{array}$$

is a pushout; equivalently, by 1.4.3, $F(A)$ is the union of $A^{\overline{0}}$ and $\overline{0}$ in $\mathrm{Sub}(1)$. So the distributive law (1.4.2), plus the fact that $(-)^{\overline{0}}$ is a right adjoint, ensures that F preserves finite intersections as a functor $\mathrm{Sub}(\overline{0}) \to \mathrm{Sub}(1)$, and hence that it is cartesian as a functor $\mathcal{P} \to \mathcal{Q}$. Note also that, since F takes values in the lattice of subterminal objects of $\mathcal{Q}$, we may identify $\mathbf{Gl}(F)$ with a full subcategory of $\mathcal{P} \times \mathcal{Q}$, consisting of those objects (A, B) such that the support of B (the image of the unique morphism $B \to 1$) is contained in $F(A)$.

Now let C be an arbitrary object of $\mathcal{E}$. We define the *thin part* C_t and the *solidification* C_s of C to be the product $C \times \overline{0}$ and the coproduct $C + \overline{0}$; clearly they are respectively thin and solid. We have a morphism

$$C_s = C + \overline{0} \xrightarrow{\eta_C + 1} (C \times \overline{0})^{\overline{0}} + \overline{0} = F(C_t)$$

(where η is the unit of $((-) \times \overline{0} \dashv (-)^{\overline{0}})$), so the pair (C_t, C_s) is an object of $\mathbf{Gl}(F)$; and the assignment $C \mapsto (C_t, C_s)$ is clearly functorial. In the other direction, we have a functor $\mathbf{Gl}(F) \to \mathcal{E}$ sending (A, B) to $A^{\overline{0}} \times B$. Now if (A, B) is an object of $\mathbf{Gl}(F)$, we have

$$(A^{\overline{0}} \times B)_t = A^{\overline{0}} \times B \times \overline{0} \cong A^{\overline{0}} \times \overline{0} \cong A$$

using the fact that B is solid and A is thin; and

$$(A^{\overline{0}} \times B)_s = (A^{\overline{0}} \times B) + \overline{0} \cong (A^{\overline{0}} \times B) + (\overline{0} \times B) \cong (A^{\overline{0}} + \overline{0}) \times B = F(A) \times B \cong B$$

where the last step uses the fact that $F(A)$ contains the support of B. Similarly, for any object C of $\mathcal{E}$, we have

$$\begin{aligned} C_t^{\overline{0}} \times C_s = C_t^{\overline{0}} \times (C + \overline{0}) &\cong (C_t^{\overline{0}} \times C) + (C_t^{\overline{0}} \times \overline{0}) \cong (C_t^{\overline{0}} \times C) + C_t \\ &= (C_t^{\overline{0}} \times C) + (\overline{0} \times C) \cong (C_t^{\overline{0}} + \overline{0}) \times C = F(C_t) \times C \cong C \end{aligned}$$

where the last step uses the fact that $F(C_t)$ contains the support of C_s, and hence also of C. The above isomorphisms are clearly natural in A, B and C; so we have an equivalence of categories, as claimed. □

Remark 2.6.8 If a quasitopos $\mathcal{E}$ is effective as a regular category, then it is easy to see that the same property is inherited by its solid part; hence, by 1.4.9 and 2.6.3(ii), the latter is a topos. It is also easy to verify that effectivity is preserved by the glueing construction; thus, as a special case of 2.6.7, we may deduce that *effective quasitoposes are exactly the categories obtained by glueing along cartesian functors from Heyting algebras to toposes.* In this sense, they form a 'minimal common generalization' of toposes and Heyting algebras.

In Example 2.6.4(d), we noted that the quasitopos $\mathbf{Sep}(\mathcal{C}, T)$ contains as a full subcategory a topos $\mathbf{Sh}(\mathcal{C}, T)$, whose subobject classifier provides the weak subobject classifier of the quasitopos. This state of affairs is typical, as we shall see shortly. First we note

Lemma 2.6.9 *Suppose $f\colon A \rightarrowtail\!\!\!\rightarrow B$ is both monic and epic in a quasitopos. Then pullback along f induces a bijection between (isomorphism classes of) cocovers $B' \rightarrowtail B$ and cocovers $A' \rightarrowtail A$.*

Proof Since isomorphism classes of cocovers $A' \rightarrowtail A$ correspond to morphisms $A \to \Omega$, the fact that f is epic ensures that pullback along it is an injective operation on cocovers. To prove that it is surjective, let $A' \rightarrowtail A$ be a cocover, and form the epimorphism-cocover factorization $A' \rightarrowtail\!\!\!\rightarrow B' \rightarrowtail B$ of the composite $A' \rightarrowtail A \rightarrowtail\!\!\!\rightarrow B$. We must show that the square

$$\begin{tikzcd} A' \arrow[r] \arrow[d, tail] & B' \arrow[d, tail] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

is a pullback. Let A'' be the pullback of B' along f; since the square above commutes, we certainly have a morphism $A' \to A''$, which is a cocover because $A' \rightarrowtail A$ is. But it is also epic, because it is the pullback of $A' \rightarrowtail\!\!\!\rightarrow B'$ along $A'' \to B'$; hence it is an isomorphism, and the square above is a pullback. □

Corollary 2.6.10 *A morphism $f\colon A \to B$ in a quasitopos is both monic and epic iff $Pf\colon PB \to PA$ is an isomorphism.*

Proof If f is both monic and epic, then so is $1_C \times f\colon C \times A \to C \times B$ for any C (since monomorphisms and epimorphisms are stable under pullback). By 2.6.9, this implies that any morphism $C \to PA$ factors uniquely through Pf, so Pf must be an isomorphism. The converse holds because P is faithful, as mentioned earlier. □

We define an object C of a quasitopos $\mathcal{E}$ to be *coarse* if 'it cannot detect the lack of balance of $\mathcal{E}$'; i.e. if, whenever $f\colon A \to B$ is both monic and epic, each morphism $A \to C$ factors uniquely through f. Thus Lemma 2.6.9 says that Ω is

coarse. Every coarse object is solid, because $0 \to \overline{0}$ is monic and epic, and the full subcategory $\mathbf{Cs}(\mathcal{E})$ of coarse objects is (trivially) closed under finite limits in $\mathcal{E}$. It is also an exponential ideal, because functors of the form $(-) \times C$ preserve monics and epics. Thus, to complete the proof that $\mathbf{Cs}(\mathcal{E})$ is a topos, it suffices to show

Lemma 2.6.11 *Let $f \colon A' \rightarrowtail A$ be a monomorphism in a quasitopos, where A is coarse. Then A' is coarse iff f is a cocover.*

Proof First suppose f is a cocover. Then, given

$$\begin{tikzcd} B' \arrow[r, "h"] \arrow[d, "g"] & A' \arrow[d, tail, "f"] \\ B & A \end{tikzcd}$$

where g is monic and epic, we get a factorization of fh through g by coarseness of A, and then we get a factorization of h through g by (the dual of) Lemma 1.3.2(ii). Conversely, suppose A' is coarse, and form the epic-cocover factorization $A' \rightarrowtail\!\!\!\!\twoheadrightarrow A'' \rightarrowtail A$ of f. By coarseness, $1_{A'}$ factors through $A' \rightarrowtail\!\!\!\!\twoheadrightarrow A''$, i.e. $A' \rightarrowtail\!\!\!\!\twoheadrightarrow A''$ is split monic; but it is also epic, and hence an isomorphism. □

Proposition 2.6.12 *In a quasitopos $\mathcal{E}$, the full subcategory $\mathbf{Cs}(\mathcal{E})$ of coarse objects is a topos, and the inclusion $\mathbf{Cs}(\mathcal{E}) \to \mathcal{E}$ has a left adjoint.*

Proof That $\mathbf{Cs}(\mathcal{E})$ is a topos, with Ω as its subobject classifier, follows from 2.6.9, 2.6.11 and the discussion preceding the latter. To obtain the reflection in $\mathbf{Cs}(\mathcal{E})$ of an arbitrary object A, form the epic-cocover factorization $A \rightarrowtail\!\!\!\!\twoheadrightarrow \overline{A} \rightarrowtail PA$ of the singleton morphism $\{\} \colon A \rightarrowtail PA$. PA is coarse by 2.6.9 and the fact that the coarse objects form an exponential ideal, so $\overline{A}$ is coarse by 2.6.11; but $A \rightarrowtail\!\!\!\!\twoheadrightarrow \overline{A}$ is monic as well as epic, because $\{\}$ is monic, and so by definition every morphism from A to a coarse object factors uniquely through it. □

In the case when $\mathcal{E}$ is the quasitopos $\mathbf{Sep}(\mathcal{C}, T)$ for a site $(\mathcal{C}, T)$, it is not hard to verify that $\mathbf{Cs}(\mathcal{E})$ is the topos $\mathbf{Sh}(\mathcal{C}, T)$. (We shall do this, in a more general context, in Section A4.4.) For $\mathcal{E} = \mathbf{Choq}$ or $\mathbf{Fre}$, the coarse objects are just the indiscrete spaces, so that $\mathbf{Cs}(\mathcal{E})$ is isomorphic to $\mathbf{Set}$; and similarly in Example 2.6.4(e) we find that $\mathbf{Cs}(\mathbf{Fuz}_f(L)) \cong \mathbf{Set}_f$. (Fuzzy-set theorists are accustomed to call these objects *crisp* rather than coarse.)

Suggestions for further reading: Adámek et al. [22], Dubuc [297], Höhle [454], Johnstone [507], Monro [861], Penon [954, 956], Wyler [1247, 1248].

A3

ALLEGORIES

A3.1 Relations in regular categories

The object of this Chapter is to present an alternative approach to the notion of topos, different from (but not entirely independent of) the one which we followed in Chapters A1 and A2, via categories of relations. While this approach will not play a very important rôle in our subsequent exposition of topos theory, it is of some interest in its own right, and it certainly provides an appealing route to the construction of some of the leading examples of toposes, which can make their structure easier to understand. The advantages of this approach have been forcefully argued over some years by P. Freyd, and a fuller account of it will be found in [381]. There are a number of other approaches to the study of categories of relations – notably that developed by A. Carboni and R. Walters [232], which makes heavier use of 2-categorical ideas – but we shall not describe them here.

By a *relation* from A to B in a cartesian category $\mathcal{C}$, we mean a subobject of $A \times B$. (In this context – in order to define an associative composition of relations – it is essential to interpret 'subobject' as meaning an isomorphism class of monomorphisms, rather than an individual monomorphism; but as usual we shall not labour the point.) We shall use the notation $A \nrightarrow B$ for a relation from A to B, and we shall generally use Greek letters to denote relations. If $(f\colon T \to A, g\colon T \to B)$ is a pair of morphisms such that the induced morphism $(f,g)\colon T \to A \times B$ is monic and belongs to the isomorphism class $\phi\colon A \nrightarrow B$, we shall call (f,g) a *tabulation* of ϕ.

The composite of two relations $\phi\colon A \nrightarrow B$ and $\psi\colon B \nrightarrow C$ can be defined in any cartesian category with images: having chosen tabulations (f,g) and (h,k) of ϕ and ψ, we form the pullback

$$\begin{tikzcd}
P \arrow[r, "q"] \arrow[d, "p"] & U \arrow[d, "h"] \\
T \arrow[r, "g"] & B
\end{tikzcd}$$

and define $\psi\phi$ to be the relation tabulated by the image of $(fp, kq)\colon P \to A \times C$. This is clearly independent of the choice of tabulations of ϕ and ψ.

Lemma 3.1.1 *Let $\mathcal{C}$ be a cartesian category with images. The composition of relations just defined is associative iff $\mathcal{C}$ is regular.*

Proof First suppose $\mathcal{C}$ is regular; let $\chi\colon C \nrightarrow D$ be a third relation, tabulated by $(l, m)\colon V \rightarrowtail C \times D$. Form the pullbacks

$$\begin{tikzcd} R \arrow[r, "u"] \arrow[d, "t"] & Q \arrow[r, "s"] \arrow[d, "r"] & V \arrow[d, "l"] \\ P \arrow[r, "q"] \arrow[d, "p"] & U \arrow[r, "k"] \arrow[d, "h"] & C \\ T \arrow[r, "g"] & B & \end{tikzcd} .$$

We claim that $\chi(\psi\phi)$ is tabulated by the image of $(fpt, msu)\colon R \to A \times D$; by symmetry $(\chi\psi)\phi$ is also tabulated by this pair, so they are equal.

To prove the claim, let

$$P \overset{e}{\twoheadrightarrow} I \overset{(a,c)}{\rightarrowtail} A \times C$$

be the image factorization of (fp, kq), and form the pullback

$$\begin{tikzcd} J \arrow[r, "y"] \arrow[d, "x"] & V \arrow[d, "l"] \\ I \arrow[r, "c"] & C \end{tikzcd} .$$

Now there exists a pullback square

$$\begin{tikzcd} R \arrow[r, "d"] \arrow[d, "t"] & J \arrow[d, "x"] \\ P \arrow[r, two heads, "e"] & I \end{tikzcd}$$

and by regularity d is a cover; so, since

$$\begin{tikzcd} R \arrow[r, two heads, "d"] \arrow[dr, "{(fpt, msu)}"'] & J \arrow[d, "{(ax, my)}"] \\ & A \times D \end{tikzcd}$$

commutes, the images of the two maps into $A \times D$ coincide. But the image of (ax, my) is by definition a tabulation of $\chi(\psi\phi)$.

Conversely, suppose composition of relations is associative: let $e: A \twoheadrightarrow B$ be a cover in $\mathcal{C}$, and $f: C \to B$ an arbitrary morphism. Let $\phi: C \nrightarrow B$, $\psi: B \nrightarrow A$ and $\chi: A \nrightarrow 1$ be tabulated by $(1_C, f)$, $(e, 1_A)$ and $(1_A, A)$ respectively (recall our convention that A also denotes the unique morphism $A \to 1$); then since $(e, A): A \to B \times 1 \cong B$ is a cover the composite $\chi\psi$ is tabulated by $(1_B, B)$, and $(\chi\psi)\phi$ is tabulated by $(1_C, C)$. But a similar calculation shows that $\chi(\psi\phi)$ is tabulated by the image of $(h, P): P \to C \times 1$, where

$$\begin{tikzcd} P \arrow[r, "h"] \arrow[d, "g"] & C \arrow[d, "f"] \\ A \arrow[r, two heads, "e"] & B \end{tikzcd}$$

is a pullback; so if these two relations are equal then h is a cover, and thus covers in $\mathcal{C}$ are stable under pullback. $\square$

Corollary 3.1.2 *If $\mathcal{C}$ is a regular category, then relations in $\mathcal{C}$ form the morphisms of a category* $\mathbf{Rel}(\mathcal{C})$ *having the same objects as $\mathcal{C}$.*

Proof Having defined composition of relations and verified that it is associative, it remains to establish the existence of identities. But it is trivial to verify that the relation ι_A tabulated by the diagonal map $(1_A, 1_A): A \rightarrowtail A \times A$ serves as the identity relation on A. $\square$

Of course, $\mathbf{Rel}(\mathcal{C})$ is more than just a category. Since the relations $A \nrightarrow B$ are the subobjects of a fixed object of $\mathcal{C}$, they have a natural partial ordering (in fact a meet-semilattice ordering, since $\mathcal{C}$ is cartesian), and this ordering is preserved under composition (that is, $\phi \leq \psi$ implies $\phi\chi \leq \psi\chi$ and $\theta\phi \leq \theta\psi$ whenever the composites are defined – although we do not in general have $(\phi \cap \psi)\chi = \phi\chi \cap \psi\chi$, since the operation of taking images does not commute with intersection). Thus $\mathbf{Rel}(\mathcal{C})$ is a particular case of a 2-category (for which see Section B1.1): in fact a *locally ordered* 2-category, i.e. one in which there is at most one 2-cell between any two given 1-cells, and the only invertible 2-cells are identities.

We shall not need to use very much 2-categorical machinery in this chapter (and in particular, we shall not have to worry about diagrams commuting only up to 2-isomorphism), but we shall need the notion of adjointness of 1-cells in a 2-category. What this amounts to in $\mathbf{Rel}(\mathcal{C})$ is that $\phi\colon A \looparrowright B$ is left adjoint to $\psi\colon B \looparrowright A$ iff $\phi\psi \leq \iota_B$ and $\iota_A \leq \psi\phi$ (the triangular identities being vacuous).

The other structure which $\mathbf{Rel}(\mathcal{C})$ possesses is an anti-involution: if $\phi\colon A \looparrowright B$ is a relation tabulated by (f,g), the *opposite relation* $\phi^\circ\colon B \looparrowright A$ is that tabulated by (g,f). It is elementary to verify that the mapping $\phi \mapsto \phi^\circ$ defines a 2-functor $\mathbf{Rel}(\mathcal{C}) \to \mathbf{Rel}(\mathcal{C})^{\mathrm{op}}$ (where the superscript 'op' signifies that 1-cells are reversed but 2-cells are not), and that its square is the identity.

If $f\colon A \to B$ is any morphism of $\mathcal{C}$, then the *graph* of f, i.e. the pair $(1_A, f)\colon A \to A \times B$, is monic (we have already used this in the proof of 3.1.1), and so tabulates a relation which we denote $f_\bullet\colon A \looparrowright B$. The assignment $f \mapsto f_\bullet$, plus the identity map on objects, defines a faithful functor $\mathcal{C} \to \mathbf{Rel}(\mathcal{C})$ (for its faithfulness, cf. the proof of 2.2.3). We shall write $f^\bullet\colon B \looparrowright A$ for $(f_\bullet)^\circ$.

Proposition 3.1.3

(i) *For any morphism f of $\mathcal{C}$, $f_\bullet$ is left adjoint to $f^\bullet$ in* $\mathbf{Rel}(\mathcal{C})$.

(ii) *Let $\phi\colon A \looparrowright B$ be a morphism in* $\mathbf{Rel}(\mathcal{C})$ *having a right adjoint. Then there is a unique morphism $f\colon A \to B$ in $\mathcal{C}$ such that $\phi = f_\bullet$.*

Proof (i) $f^\bullet f_\bullet$ is tabulated by the kernel-pair of f, and the reflexiveness of this relation (cf. 1.3.6) is precisely the inequality $\iota_A \leq f^\bullet f_\bullet$. On the other hand, $f_\bullet f^\bullet$ is the image of $(f,f)\colon A \to B \times B$, and since this morphism factors through the diagonal we have $f_\bullet f^\bullet \leq \iota_B$.

(ii) Let ψ be the right adjoint of ϕ, and let $(a,b)\colon T \rightarrowtail A \times B$, $(b',a')\colon T' \rightarrowtail B \times A$ be tabulations of ϕ and ψ. First, the inequality $\iota_A \leq \psi\phi$ says that the image of $(ap, a'p')\colon P \to A \times A$ contains the diagonal, where

$$\begin{tikzcd} P \arrow[r, "p"] \arrow[d, "p'"] & T \arrow[d, "b"] \\ T' \arrow[r, "b'"] & B \end{tikzcd}$$

is a pullback. So ap factors as a cover followed by a split epimorphism; hence it is a cover, and in particular a is a cover. Similarly, a' is a cover, so in the pullback

$$\begin{tikzcd} Q \arrow[r, "q"] \arrow[d, "q'"] & T \arrow[d, "a"] \\ T' \arrow[r, "a'"] & A \end{tikzcd}$$

both q and q' are covers. But the inequality $\phi\psi \leq \iota_B$ means that $bq = b'q'$; and since we also have $aq = a'q'$ this forces $\psi = \phi^\circ$, since each relation is tabulated by the image of $(bq, aq) = (b'q', a'q')$. Thus we can now drop the primes on T', a' and b'. Also, since we have $aq = aq'$ and $bq = bq'$ and the pair (a, b) is monic, we must have $q = q'$. But (q, q') is the kernel-pair of a; so a is monic, and hence an isomorphism. Thus, putting $f = ba^{-1}$, we have $\phi = f_\bullet$, since a itself is an isomorphism from the given tabulation of ϕ to the graph of f. The uniqueness of f follows from the faithfulness of $f \mapsto f_\bullet$. □

Thus the regular category $\mathcal{C}$ is recoverable (up to isomorphism) from the 2-category $\mathbf{Rel}(\mathcal{C})$, as the subcategory containing all the objects of $\mathbf{Rel}(\mathcal{C})$ but only those morphisms which have right adjoints. Because of this, a morphism in a locally ordered 2-category which has a right adjoint is often called a *map*; of course, such morphisms always form a subcategory. In the next section, we shall consider how to characterize categories of the form $\mathbf{Rel}(\mathcal{C})$ amongst locally ordered 2-categories.

In passing, we note

Scholium 3.1.4 *Let $f: A \to B$ be a morphism in a regular category.*

(i) *f is monic iff the unit $\iota_A \leq f^\bullet f_\bullet$ is an identity.*

(ii) *f is a cover iff the counit $f_\bullet f^\bullet \leq \iota_B$ is an identity.*

Proof Both observations are immediate from the construction of these inequalities in the proof of 3.1.3(i). □

We have already remarked that composition of relations does not in general distribute over the finite intersections which exist in the 'hom-posets' of $\mathbf{Rel}(\mathcal{C})$. The following result, known as the *modular law* (for reasons which will become clear in 3.2.2(a) below), provides a substitute in some cases for the missing distributive law.

Proposition 3.1.5 *Let $\phi: A \nrightarrow B$, $\psi: B \nrightarrow C$ and $\chi: A \nrightarrow C$ be three relations in a regular category. Then*

$$\psi\phi \cap \chi \leq (\psi \cap \chi\phi^\circ)\phi \, .$$

Proof Let $(f, g): T \rightarrowtail A \times B$, $(h, k): U \rightarrowtail B \times C$ and $(l, m): V \rightarrowtail A \times C$ be tabulations of the three relations. First form the limit of the diagram

$$\begin{tikzcd} T \arrow[r, "f"] \arrow[d, "g"'] & A & V \arrow[l, "l"'] \arrow[d, "m"] \\ B & U \arrow[l, "h"'] \arrow[r, "k"] & C \end{tikzcd} \quad ;$$

by definition, this is an object P with morphisms $x: P \to T$, $y: P \to U$ and $z: P \to V$ satisfying $fx = lz$, $gx = hy$ and $ky = mz$ (and universal among such). (If we were working with relations in the category **Set**, P would be the set of all triples (a, b, c) such that a is ϕ-related to b, b is ψ-related to c and a is χ-related to c. The underlying idea of this proof is that if (a, c) is in (a tabulation of) $\psi\phi \cap \chi$, then there must exist b such that (a, b, c) is in P; but then (b, c) is in $\psi \cap \chi\phi^\circ$ and (a, b) is in ϕ.)

We claim first that the image of $(lz, mz): P \to A \times C$ is a tabulation of $\psi\phi \cap \chi$. To prove this claim, note that a tabulation of $\psi\phi \cap \chi$ is obtained by constructing the pullback

$$\begin{tikzcd} Q \arrow[r, "t"] \arrow[d, "u"'] & T \arrow[d, "g"] \\ U \arrow[r, "h"] & B \end{tikzcd} ,$$

the image factorization $Q \twoheadrightarrow I \rightarrowtail A \times C$ of (ft, ku), and then the pullback

$$\begin{tikzcd} R \arrow[r, tail] \arrow[d, tail] & I \arrow[d, tail] \\ V \arrow[r, tail, "{(l,m)}"] & A \times C \end{tikzcd} .$$

But an elementary diagram-chase shows that there is a pullback

$$\begin{tikzcd} P \arrow[r] \arrow[d, "z"'] & Q \arrow[d, "{(ft,ku)}"] \\ V \arrow[r, "{(l,m)}"] & A \times C \end{tikzcd} ,$$

so since image factorizations are stable under pullback there is a cover $P \twoheadrightarrow R$.

Similarly, the image of $(hy, ky): P \to B \times C$ (say $S \rightarrowtail B \times C$) is a tabulation of $\psi \cap \chi\phi^\circ$. Now the cover $P \twoheadrightarrow S$, together with the morphism $x: P \to T$, induce a morphism from P to the pullback

$$\begin{tikzcd} W \arrow[r] \arrow[d] & S \arrow[d] \\ T \arrow[r, "g"] & B \end{tikzcd} ;$$

this is easily seen to be a morphism over $A \times C$, and hence to induce an inequality between the images of the two morphisms into $A \times C$. But the image of $W \rightarrowtail T \times S \to A \times C$ is by definition a tabulation of $(\psi \cap \chi\phi^\circ)\phi$. □

Corollary 3.1.6 *Let ϕ be a morphism of* $\mathbf{Rel}(\mathcal{C})$ *such that $\phi\phi^\circ \leq \iota_B$ (where B is the codomain of ϕ). Then the distributive law*

$$(\psi \cap \chi)\phi = \psi\phi \cap \chi\phi$$

holds whenever the composites are defined.

Proof We have $(\psi \cap \chi)\phi \geq (\psi \cap \chi\phi\phi^\circ)\phi \geq \psi\phi \cap \chi\phi$ by the given assumption and the modular law, but the reverse inequality always holds because $(-)\phi$ is order-preserving. □

The inequality $\phi\phi^\circ \leq \iota_B$ holds in particular when ϕ is a map, by 3.1.3; but in this case it would be easy to prove the distributive law directly.

Suggestions for further reading: Carboni [218], Carboni & Walters [232], Freyd & Scedrov [381], Kelly [584].

A3.2 Allegories and tabulations

Abstracting from the categories of relations considered in the last section, we define an *allegory* to be a locally ordered 2-category $\mathcal{A}$ whose hom-posets have binary intersections (note that we don't, for the moment, require them to have top elements), equipped with an anti-involution $\phi \mapsto \phi^\circ$ and satisfying the modular law

$$\psi\phi \cap \chi \leq (\psi \cap \chi\phi^\circ)\phi$$

whenever this makes sense. As is already apparent, we shall maintain continuity with the last section by using Greek letters (and the notation $\phi\colon A \looparrowright B$) for arbitrary morphisms in an allegory, reserving italic letters and ordinary arrows for maps (i.e. morphisms with right adjoints). We write $\mathbf{Map}(\mathcal{A})$ for the subcategory of maps of $\mathcal{A}$; note that it contains all isomorphisms of $\mathcal{A}$ (and they are isomorphisms in $\mathbf{Map}(\mathcal{A})$), since an isomorphism is both left and right adjoint to its inverse. (Partly for this reason, we shall denote identity morphisms in $\mathcal{A}$ by 1 rather than ι.)

Because of the anti-involution, the apparent asymmetry in the modular law is only apparent; it implies (indeed, is equivalent to) the law

$$\psi\phi \cap \chi \leq \psi(\phi \cap \psi^\circ\chi)\ .$$

A further immediate deduction from the modular law, which we note for future reference, is

Lemma 3.2.1 *For any morphism ϕ in an allegory, we have $\phi \leq \phi\phi^\circ\phi$.*

Proof We have $\phi = 1\phi \cap \phi \leq (1 \cap \phi\phi^\circ)\phi \leq \phi\phi^\circ\phi$. □

Before proceeding further, we arm ourselves with a couple of new examples.

Examples 3.2.2 (*a*) Let L be a lattice. We may regard L as a locally ordered 2-category with one object, interpreting composition as join in L (and intersection as meet in L); since join is commutative, the identity mapping $\lambda^\circ = \lambda$ is an anti-involution. In this context, the modular law is equivalent to the law familiar to lattice-theorists under that name, i.e.

$$\lambda\mu \cap \nu \leq \lambda(\mu \cap \nu) \quad \text{whenever} \quad \lambda \leq \nu;$$

the point is that the latter inequality is equivalent to $\nu = \nu\lambda^\circ$. Note that any lattice satisfies $\lambda \leq \lambda\lambda^\circ\lambda$ (in fact we have equality here), so this inequality does not imply the modular law. In fact, as is well known, any lattice generated by two elements is modular (and even distributive), so the modular law cannot be equivalent to any collection of identities in two or fewer variables.

(*b*) Let L be a distributive lattice. The allegory $\mathbf{Mat}_f(L)$ of finite L-valued matrices has finite sets (or natural numbers, if you prefer) as objects, and morphisms $A \looparrowright B$ are '$B \times A$ matrices with entries in L', i.e. functions $B \times A \to L$, the partial ordering (and binary intersection) on hom-sets being computed pointwise. Composition is defined by matrix multiplication, as in 1.4.6: the composite of $\phi\colon A \looparrowright B$ and $\psi\colon B \looparrowright C$ is the function χ defined by

$$\chi(c, a) = \bigvee\{\psi(c, b) \wedge \phi(b, a) \mid b \in B\},$$

the distributive law in L being precisely what is needed to verify that this composition is associative. The identity matrix 1_A is what you think it is, and ϕ° is simply the transpose of ϕ. The verification of the modular law is straightforward. Moreover, if L is a frame (i.e. L has arbitrary joins, and binary meets distribute over them), we may embed this example in a larger allegory $\mathbf{Mat}(L)$ whose objects are arbitrary sets, the rest of the definition remaining unchanged.

If L is the two-element lattice $\mathbf{2} = \{0, 1\}$, then $\mathbf{Mat}_f(L)$ is simply another description of the allegory $\mathbf{Rel}(\mathbf{Set}_f)$, using the bijection between subsets of $A \times B$ and their characteristic functions. For a more general L, however, it is instructive to compute the maps $A \to B$ in $\mathbf{Mat}_f(L)$ (or in $\mathbf{Mat}(L)$): using 3.2.3 below, they are easily shown to be those matrices whose columns are disjoint decompositions of 1 in L, i.e. those ϕ such that $\phi(b, a) \wedge \phi(b', a) = 0$ whenever $b \neq b'$, and

$$\bigvee\{\phi(b, a) \mid b \in B\} = 1$$

for each $a \in A$. Thus we see that $\mathbf{Map}(\mathbf{Mat}_f(L))$ is equivalent to a full subcategory of the category $\mathbf{Pos}(L)$ of 1.4.6, consisting of those objects which are copowers of 1 in $\mathbf{Pos}(L)$. (We shall see how to recover the 'missing objects',

by splitting cores in $\mathbf{Mat}_f(L)$, in 3.3.7 below.) Note in particular that if 1 has no nontrivial disjoint decompositions in L (for example, if L is totally ordered), then any map in $\mathbf{Mat}_f(L)$ has all its entries 0 or 1; thus $\mathbf{Map}(\mathbf{Mat}_f(L)) \cong \mathbf{Map}(\mathbf{Mat}_f(\mathbf{2})) \cong \mathbf{Set}_f$, although $\mathbf{Mat}_f(L)$ is clearly not equivalent to $\mathbf{Mat}_f(\mathbf{2})$ unless $L \cong \mathbf{2}$.

(c) The axioms for an allegory do not include anything (such as the existence of limits or power-objects) which allows us to construct new objects from old; it is therefore immediate that *any full subcategory of an allegory is an allegory*, no matter how bizarre its collection of objects. For example, the collection of all sets with either 23 or 35 elements, and all relations between them, is an allegory. At first sight, this may look like a defect of the allegory axioms, but it is actually one of their strengths: the point is that many allegories of interest to us – that is, allegories whose categories of maps are toposes – may be generated (by the idempotent-splitting processes described in the next section) from quite small (in the non-technical sense) full sub-allegories, and these sub-allegories are often easier to describe explicitly than the large categories in which we are interested.

Next, we prove an important result (already established for allegories of the form $\mathbf{Rel}(\mathcal{C})$ in the course of proving 3.1.3), which is the first indication we have had of the real power of the modular law.

Lemma 3.2.3

(i) *In an allegory, the right adjoint of a map f is necessarily f°.*

(ii) *The maps in an allegory are discretely ordered (i.e. $f \leq g$ implies $f = g$).*

Proof (i) Let $f\colon A \to B$ be a map. We can certainly write its right adjoint as g° where g is a map, since an adjunction $f \dashv \phi$ yields an adjunction $\phi^\circ \dashv f^\circ$. Now we have

$$(f^\circ \cap g^\circ)f = (g^\circ \cap 1_A f^\circ)f \geq g^\circ f \cap 1_A = 1_A$$

by the modular law and the fact that $g^\circ f \geq 1_A$; and clearly $f(f^\circ \cap g^\circ) \leq fg^\circ \leq 1_B$. So $f^\circ \cap g^\circ = (f \cap g)^\circ$ is also right adjoint to f, but by the uniqueness of adjoints this forces $(f \cap g)^\circ = g^\circ$, i.e. $f \cap g = g$. By symmetry, we also have $f \cap g = f$, so $f = g$.

(ii) Let f and g be maps such that $f \leq g$. Then we also have $f^\circ \leq g^\circ$, and hence

$$g \leq gf^\circ f \leq gg^\circ f \leq f,$$

so $f = g$. □

Let $\phi\colon A \looparrowright B$ be a morphism in an allegory. By a *tabulation* of ϕ we mean a pair of maps $f\colon T \to A$, $g\colon T \to B$ with common domain satisfying $\phi = gf^\circ$ and $f^\circ f \cap g^\circ g = 1_T$. It is easy to see that tabulations in $\mathbf{Rel}(\mathcal{C})$ are (modulo the isomorphism $\mathbf{Map}(\mathbf{Rel}(\mathcal{C})) \cong \mathcal{C}$) exactly what we called tabulations in the previous section; the second condition in the definition says that the pair (f, g) is jointly monic (cf. 3.1.4(ii)). Thus every morphism in $\mathbf{Rel}(\mathcal{C})$ has a tabulation; we call an allegory *tabular* if this condition holds.

Clearly, not all allegories are tabular. In a modular lattice L, regarded as an allegory as in 3.2.2(a), only the identity morphism (which is the element 0 of L!) is a map, and hence it is the only morphism with a tabulation. Again, if L is a distributive lattice with more than two elements, the allegory $\mathbf{Mat}_f(L)$ of 3.2.2(b) is not tabular: if λ is any element of L other than 0 or 1, then it is easy to verify that the matrix $\phi\colon \{a\} \looparrowright \{a\}$ defined by $\phi(a,a) = \lambda$ cannot be tabulated. However, in the next section we shall see how to embed certain allegories (including $\mathbf{Mat}_f(L)$) into tabular allegories in a canonical way.

In passing, we note that in a tabular allegory the operation $\phi \mapsto \phi^\circ$ may be derived from the other structure; for if f is a map then f° must be its right adjoint, and if $\phi = gf^\circ$ then $\phi^\circ = fg^\circ$. In theory, therefore, one could axiomatize the notion of tabular allegory without mentioning this operation; but in practice it would make the statement of the modular law intolerably opaque.

Lemma 3.2.4 *Suppose $\phi\colon A \looparrowright B$ has a tabulation (f,g), and let $x\colon C \to A$, $y\colon C \to B$ be maps. Then $yx^\circ \le \phi$ iff there exists a map h such that $x = fh$ and $y = gh$. Moreover, if such an h exists it is unique.*

Proof If $x = fh$ and $y = gh$, then $yx^\circ = ghh^\circ f^\circ \le gf^\circ = \phi$ since h is a map. Conversely, suppose $yx^\circ \le \phi$; define $\theta = f^\circ x \cap g^\circ y\colon C \looparrowright T$. Since x and y are maps, we have $1_C \le y^\circ y x^\circ x \le y^\circ g f^\circ x$. Now

$$\begin{aligned}\theta^\circ\theta &= (x^\circ f \cap y^\circ g)(f^\circ x \cap (f^\circ x \cap g^\circ y))\\ &\ge (x^\circ f \cap y^\circ g) f^\circ x \cap 1_C\\ &\ge 1_C \cap y^\circ g f^\circ x \cap 1_C = 1_C\end{aligned}$$

by two applications of the modular law, and

$$\begin{aligned}\theta\theta^\circ &= (f^\circ x \cap g^\circ y)(x^\circ f \cap y^\circ g)\\ &\le f^\circ x x^\circ f \cap g^\circ y y^\circ g\\ &\le f^\circ f \cap g^\circ g = 1_T,\end{aligned}$$

so θ is a map (and we can now denote it by h). Next,

$$fh = f(f^\circ x \cap g^\circ y) \le f f^\circ x \le x$$

whence $fh = x$ by 3.2.3(ii), and similarly $gh = y$. Finally, if k is any other map satisfying $fk = x$ and $gk = y$, then

$$k = (f^\circ f \cap g^\circ g)k \le f^\circ f k \cap g^\circ g k = f^\circ x \cap g^\circ y = h,$$

whence $h = k$ by 3.2.3(ii) again. □

Corollary 3.2.5 *Any two tabulations of a given morphism in an allegory are uniquely isomorphic.*

Proof Apply 3.2.4 each way round to the two tabulations. □

Lemma 3.2.6 *Let $f: A \to B$ be a map in a tabular allegory $\mathcal{A}$. The following are equivalent:*

(i) *f is monic in* $\mathbf{Map}(\mathcal{A})$.

(ii) $f^\circ f = 1_A$.

(iii) *(f, f) is a tabulation of ff°.*

Proof (i) $\Rightarrow$ (ii): Suppose f is monic, and let (h, k) be a tabulation of $f^\circ f$. From the equality $f^\circ f = kh^\circ$ we deduce $fk \leq fkh^\circ h = ff^\circ fh = fh$, whence $fk = fh$ by 3.2.3(ii), whence $k = h$ since f is monic. So $1_A \leq f^\circ f = hh^\circ \leq 1_A$.

(ii) $\Rightarrow$ (iii) is immediate from the definition of a tabulation.

(iii) $\Rightarrow$ (i): If (f, f) is a tabulation of anything, then it follows from the uniqueness clause of 3.2.4 that f is monic in $\mathbf{Map}(\mathcal{A})$. □

Proposition 3.2.7 *Let $\mathcal{A}$ be a tabular allegory. Then* $\mathbf{Map}(\mathcal{A})$ *is locally regular (that is, it has pullbacks and images, and covers are stable under pullback).*

Proof First we observe that if f, g, h, k are four maps satisfying $hf = kg$, then $gf^\circ \leq k^\circ kgf^\circ = k^\circ hff^\circ \leq k^\circ h$; and conversely if $gf^\circ \leq k^\circ h$, then $kg \leq kgf^\circ f \leq kk^\circ hf \leq hf$, whence $hf = kg$ by 3.2.3(ii). Thus 3.2.4 tells us that we may construct a pullback for a pair of maps (h, k) with common codomain by finding a tabulation of $k^\circ h$. Next, to form the image of a map $f: A \to B$, consider a tabulation of ff°, say $(g: T \to B, h: T \to B)$. Since $ff^\circ \leq 1_B$ and 1_B is tabulated by $(1_B, 1_B)$, 3.2.4 yields a unique map k such that $1_B k = g = h$; that is, $g = h$. It now follows from 3.2.6 that g is monic in $\mathbf{Map}(\mathcal{A})$, and 3.2.4 yields a unique map $q: A \to T$ with $gq = f$. Moreover, if $m: C \rightarrowtail B$ is any monic map through which f factors, then $gg^\circ = ff^\circ \leq mm^\circ$; since (m, m) tabulates mm° by 3.2.6, 3.2.4 now yields $g \leq m$ in $\mathrm{Sub}(B)$. So (q, g) is an image factorization of f.

In particular, we see that f is a cover in $\mathbf{Map}(\mathcal{A})$ iff $ff^\circ = 1_B$. Now suppose that we have a pullback square

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "h"] \\ C \arrow[r, "k"] & D \end{tikzcd}$$

in which h is a cover; then $1_C \leq k^\circ k = k^\circ hh^\circ k = gf^\circ fg^\circ$ (the last step by the construction of pullbacks given earlier). So

$$1_C = 1_C \cap gf^\circ fg^\circ \leq (g \cap gf^\circ f)g^\circ = gg^\circ$$

using the modular law and the inequality $1_A \leq f^\circ f$; thus g is a cover. □

Thus we have shown that tabular allegories are essentially the same thing as categories of relations of locally regular categories. (The proof given in the last section that **Rel**$(\mathcal{C})$ is an allegory goes through almost unchanged if $\mathcal{C}$ lacks a terminal object (for example, if $\mathcal{C}$ is the category **LH** of spaces and local homeomorphisms, cf. 1.2.7), although we have to strengthen the definition of 'locally regular category' slightly. The notion of relation must of course be redefined if $\mathcal{C}$ fails to have binary products; we must think of a relation as an isomorphism class of jointly-monic pairs rather than a subobject of a product – note that in general the poset of relations from A to B will fail to have a top element. In order to construct binary intersections of relations, we need to have equalizers as well as pullbacks in $\mathcal{C}$, and in order to define composition we need images for pairs of morphisms as well as single morphisms – i.e. we need to know that each pair of morphisms with common domain factors in a best possible way through a jointly-monic pair. In fact it is not hard to verify, by arguments similar to those in the proof of 3.2.7, that both these extra conditions are satisfied by **Map**$(\mathcal{A})$ for any tabular allegory $\mathcal{A}$; they are also satisfied by **LH**.)

To characterize the categories **Rel**$(\mathcal{C})$ where $\mathcal{C}$ is regular, we merely have to add some condition to the axioms for a tabular allegory $\mathcal{A}$ which ensures that **Map**$(\mathcal{A})$ has a terminal object. The following definition is only slightly less blatant than demanding this explicitly, but it seems to be the best we can do: an object U of $\mathcal{A}$ is called a *unit* if 1_U is the largest morphism $U \looparrowright U$, and for every object A there is a morphism $\phi\colon A \looparrowright U$ with $\phi^\circ\phi \geq 1_A$.

Lemma 3.2.8 *If U is a unit in an allegory $\mathcal{A}$, then it is a terminal object in* **Map**$(\mathcal{A})$. *The converse holds if $\mathcal{A}$ is tabular.*

Proof If U is a unit, then there certainly exists a map $A \to U$ for every A, since every $\phi\colon A \looparrowright U$ satisfies $\phi\phi^\circ \leq 1_U$. But if f and g are two such maps, then we have $fg^\circ \leq 1_U$, so $f \leq fg^\circ g \leq g$, whence $f = g$ by 3.2.3(ii).

Conversely, if U is terminal in **Map**$(\mathcal{A})$, then the second condition for being a unit is clearly satisfied. For the first, let $\phi\colon U \looparrowright U$; then in any tabulation (f, g) of ϕ we have $f = g$, whence $\phi = gf^\circ = ff^\circ \leq 1_U$. □

For future reference, we also note

Lemma 3.2.9 *If an allegory has a unit, then each of its hom-posets has a top element.*

Proof Let U be a unit in $\mathcal{A}$, and for each object A let p_A denote the unique map $A \to U$. If $\phi\colon A \looparrowright U$, then $\phi \leq \phi p_A^\circ p_A \leq p_A$ since $\phi p_A^\circ \leq 1_U$. Now for any $\psi\colon A \looparrowright B$ we have $p_B\psi \leq p_A$, so that $\psi \leq p_B^\circ p_A$ by adjointness, i.e. $p_B^\circ p_A$ is the top element of $\mathcal{A}(A, B)$. □

We are now ready to put everything together:

Theorem 3.2.10
(i) *If $\mathcal{C}$ is a regular category, then* **Rel**$(\mathcal{C})$ *is a tabular allegory with a unit, and* $\mathcal{C} \cong$ **Map**(**Rel**$(\mathcal{C})$).

(ii) *If $\mathcal{A}$ is a tabular allegory with a unit, then* $\mathbf{Map}(\mathcal{A})$ *is a regular category, and* $\mathcal{A} \cong \mathbf{Rel}(\mathbf{Map}(\mathcal{A}))$.

Proof The only part of this that we have not already verified is the isomorphism $\mathcal{A} \cong \mathbf{Rel}(\mathbf{Map}(\mathcal{A}))$. But the definition of tabularity and 3.2.5 ensure that we have a bijection between morphisms $A \looparrowright B$ in $\mathcal{A}$ and relations $A \looparrowright B$ in $\mathbf{Map}(\mathcal{A})$; and the construction of pullbacks and images in $\mathbf{Map}(\mathcal{A})$, given in the proof of 3.2.7, make it straightforward (though not entirely trivial) to verify that this bijection commutes with composition – the fact that it preserves the partial order on hom-sets is immediate from 3.2.4, and that it commutes with $\phi \mapsto \phi^\circ$ is trivial. □

Theorem 3.2.10 can be made functorial. If we define a (*unital*) *morphism of allegories* to be a functor commuting with intersections and with $\phi \mapsto \phi^\circ$ (and preserving the unit), then it is easy to see that unital morphisms between tabular allegories with unit correspond to regular functors between regular categories. (It is possible, though less straightforward, to make it 2-functorial; the right notion of 'natural transformation' between morphisms of allegories is not just that of a natural transformation of functors, but something more complicated. We shall not go into details here.)

It is also possible to obtain more restricted equivalences from 3.2.10, by adding further categorical or allegorical structure to each side. We give two examples of this here; further examples will occur in the next two sections.

Proposition 3.2.11 *Let $\mathcal{A}$ be a tabular allegory with a unit.*

(i) $\mathbf{Map}(\mathcal{A})$ *is a coherent category iff the hom-posets of $\mathcal{A}$ have finite unions (including least elements) and composition distributes over these.*

(ii) $\mathbf{Map}(\mathcal{A})$ *is a positive coherent category iff $\mathcal{A}$ has the structure described in (i), together with finite products (equivalently, finite coproducts).*

Proof (i) It is clear that if $\mathcal{C}$ is a coherent category then the hom-posets of $\mathbf{Rel}(\mathcal{C})$ have finite unions; and the fact that unions are stable under pullback in $\mathcal{C}$ makes it easy to verify that composition in $\mathbf{Rel}(\mathcal{C})$ distributes over unions. Conversely, if $\mathbf{Rel}(\mathcal{C})$ has the extra structure indicated, we deduce that $\mathcal{C}$ is coherent by observing that subobjects of A in $\mathcal{C}$ correspond to relations $A \looparrowright 1$, and pulling back subobjects along $f: A \to B$ corresponds to composing such relations with $f_\bullet$.

(ii) If A and B have a disjoint coproduct $A \amalg B$, then for any object C we have $C \times (A \amalg B) \cong (C \times A) \amalg (C \times B)$, and any subobject of $C \times (A \amalg B)$ decomposes uniquely as the disjoint union of a subobject of $C \times A$ and a subobject of $C \times B$. Using this, it is easy to see that $A \amalg B$ is both a product and a coproduct of A and B in $\mathbf{Rel}(\mathcal{C})$ – and the initial object of $\mathcal{C}$, being strict, is both initial and terminal in $\mathbf{Rel}(\mathcal{C})$.

Conversely, let $\mathcal{A}$ be an allegory satisfying the conditions of (i), and suppose two objects A and B have a coproduct $A + B$ in $\mathcal{A}$. (Of course, it is also their

product, because of the anti-involution.) Let $\nu_1 : A \looparrowright A + B$, $\nu_2 : B \looparrowright A + B$ be the coprojections; let $\pi_1 : A + B \looparrowright A$ be the unique morphism satisfying $\pi_1\nu_1 = 1_A$ and $\pi_1\nu_2 = 0$, and let $\pi_2 : A + B \looparrowright B$ be defined similarly. Consider the morphism $\theta = \nu_1\pi_1 \cup \nu_2\pi_2 : A + B \looparrowright A + B$; since composition distributes over unions, we have $\theta\nu_1 = \nu_1 \cup 0 = \nu_1$ and $\theta\nu_2 = 0 \cup \nu_2 = \nu_2$. So by the uniqueness of morphisms out of coproducts, we have $\theta = 1_{A+B}$; in particular, $\nu_1\pi_1 \leq 1_{A+B}$. But $\pi_1\nu_1 = 1_A$, so ν_1 is a map, and $\pi_1 = \nu_1^\circ$ by 3.2.3(i). In fact ν_1 is monic in $\mathbf{Map}(\mathcal{A})$, since it is split monic in $\mathcal{A}$. Similarly, ν_2 is a monic map; so we can regard A and B as subobjects of $A+B$ in $\mathbf{Map}(\mathcal{A})$. To compute the intersection of these subobjects, the construction of 3.2.7 tells us to tabulate the morphism $\nu_2^\circ\nu_1 = \pi_2\nu_1 = 0$; but this clearly yields the smallest subobject of $A+B$, so A and B are disjoint. It now follows from 1.4.4 that A and B have a disjoint coproduct in $\mathbf{Map}(\mathcal{A})$ (in fact, though we don't need this, the coproduct is $A + B$ itself, since the equation $\nu_1\nu_1^\circ \cup \nu_2\nu_2^\circ = 1$ indicates that the union of the two given subobjects is the whole of $A + B$). □

If $\mathcal{A}$ is an allegory (not necessarily tabular) whose hom-posets have finite unions which are distributed over by composition, we shall call it a *union allegory*. If $\mathcal{A}$ additionally has finite (co)products, we call it a *positive union allegory*. Part of the preceding proof can now be recast:

Scholium 3.2.12 *If two objects of a union allegory $\mathcal{A}$ have a coproduct, it is also their coproduct in* $\mathbf{Map}(\mathcal{A})$.

Proof Suppose the coproduct $A+B$ exists in $\mathcal{A}$. The argument in the proof of 3.2.11(ii) shows that the coprojections are maps (and we shall now denote them by $n_1 : A \to A + B$, $n_2 : B \to A + B$), so any map with domain $A + B$ induces by composition a pair of maps with domains A and B. Conversely, suppose we are given maps $f : A \to C$, $g : B \to C$; we must show that the unique morphism $\phi : A + B \looparrowright C$ induced by f and g is a map. But we have $\phi = \phi(n_1n_1^\circ \cup n_2n_2^\circ) = fn_1^\circ \cup gn_2^\circ$, so

$$\begin{aligned}\phi^\circ\phi &= (n_1f^\circ \cup n_2g^\circ)(fn_1^\circ \cup gn_2^\circ)\\ &\geq n_1f^\circ fn_1^\circ \cup n_2g^\circ gn_2^\circ\\ &\geq n_1n_1^\circ \cup n_2n_2^\circ = 1_{A+B},\end{aligned}$$

and

$$\begin{aligned}\phi\phi^\circ &= (fn_1^\circ \cup gn_2^\circ)(n_1f^\circ \cup n_2g^\circ)\\ &= fn_1^\circ n_1f^\circ \cup fn_1^\circ n_2g^\circ \cup gn_2^\circ n_1f^\circ \cup gn_2^\circ n_2g^\circ\\ &= ff^\circ \cup 0 \cup 0 \cup gg^\circ \leq 1_C\,.\end{aligned}$$

□

A similar argument shows that the initial object 0, if it exists, in a union allegory $\mathcal{A}$ is also initial in $\mathbf{Map}(\mathcal{A})$ (in fact strict initial, since the unique morphism $A \looparrowright 0$, being the opposite of a map, can only be a map if it is an isomorphism). The argument also extends readily to infinite coproducts, provided the hom-posets of $\mathcal{A}$ have infinite unions and composition distributes over them. In the

context of a particular model of set theory, we shall call an allegory *geometric* if its hom-posets are complete lattices (i.e. small posets with arbitrary unions) and composition distributes over arbitrary unions, and *positive geometric* if it additionally has arbitrary set-indexed coproducts. Of course, geometric allegories which are tabular and have units correspond to the geometric categories introduced at the end of section A1.4 (and positive geometric allegories similarly correspond to ∞-pretoposes, provided their equivalences split – cf. 3.3.9 below.)

Example 3.2.13 For a distributive lattice L, $\mathbf{Mat}_f(L)$ is a positive union allegory. Unions of matrices are computed pointwise, and it is easy to verify that composition distributes over them; the coproduct of two objects A and B is just their disjoint union, since a matrix $\phi\colon C \times (A \amalg B) \to L$ is uniquely determined by the pair $(\phi|_{C\times A}, \phi|_{C\times B})$. Similarly, if L is a frame, then $\mathbf{Mat}(L)$ is a positive geometric allegory.

It should be noted that, in our definition of union allegory (and of geometric allegory), we have not required that intersection of morphisms should distribute over union. For tabular union allegories, this distributive law automatically holds, by 3.2.11 and 1.4.2, but in general it may fail: a counterexample is provided by any modular lattice which is not distributive, regarded as an allegory as in 3.2.2(a). Freyd and Scedrov [381] use the term 'distributive allegory' for a union allegory in which intersections distribute over unions.

Suggestion for further reading: Freyd & Scedrov [381]

A3.3 Splitting symmetric idempotents

In this section we shall be concerned with various classes of endomorphisms in an allegory. We begin with a clutch of definitions.

Definition 3.3.1 Let $\phi\colon A \looparrowright A$ be an endomorphism in an allegory.

(a) We say ϕ is *reflexive* if $1_A \leq \phi$.

(b) We say ϕ is *symmetric* if $\phi \leq \phi^\circ$ (equivalently, $\phi = \phi^\circ$).

(c) We say ϕ is *transitive* if $\phi\phi \leq \phi$.

(d) We say ϕ is an *equivalence* if it has all three of the above properties.

(e) We say ϕ is a *core* if $\phi \leq 1_A$. (The name 'core' is short for 'coreflexive'; but there is no connection with the coreflexive pairs considered in 1.2.9.)

It is clear that, in an allegory of the form $\mathbf{Rel}(\mathcal{C})$, the first three properties are equivalent to those (of tabulations of the relations) given the same names in 1.3.6. Thus equivalences in $\mathbf{Rel}(\mathcal{C})$ are the same thing as equivalence relations in $\mathcal{C}$, as defined there. (We remark that there is little danger of confusion between this use of the word 'equivalence' and its usual meaning in a 2-category (for which see Section B1.1); since the only 2-isomorphisms in an allegory are identities,

the only equivalences in the 2-categorical sense are isomorphisms, and we do not need a separate name for them.)

We also, of course, call ϕ *idempotent* if $\phi\phi = \phi$. Clearly, an equivalence (indeed, any reflexive transitive morphism) is idempotent; but we also have

Lemma 3.3.2
(i) *A symmetric transitive morphism is idempotent.*
(ii) *A core is symmetric and idempotent.*

Proof (i) From 3.2.1, we have $\phi \leq \phi\phi^\circ\phi$; so if ϕ is symmetric and transitive, we have $\phi \leq \phi\phi\phi \leq \phi\phi \leq \phi$.

(ii) The inequality $\phi \leq \phi^\circ$ also follows from $\phi \leq \phi\phi^\circ\phi$, given $\phi \leq 1_A$. And a core is trivially transitive, and hence idempotent by (i). □

We recall the construction of the category $\mathcal{C}[\check{\mathcal{E}}]$, where $\mathcal{E}$ is a class of idempotents in a category $\mathcal{C}$, described in 1.1.8. What is of interest to us now is that if $\mathcal{C}$ is an allegory and $\mathcal{E}$ consists of *symmetric* idempotents, then $\mathcal{C}[\check{\mathcal{E}}]$ also has an allegory structure. Before proving this, we need

Lemma 3.3.3 *Let $\epsilon\colon A \looparrowright A$ be a symmetric idempotent in an allegory.*

(i) *Any splitting of ϵ must have the form $\phi\phi^\circ$ for some $\phi\colon B \looparrowright A$.*
(ii) *If ϵ is a core, then any splitting of it must have the form ff° where f is a map (in fact a monomorphism in* **Map**$(\mathcal{A})$*).*
(iii) *If ϵ is an equivalence, then any splitting of it must have the form $g^\circ g$ where g is an epimorphism in* **Map**$(\mathcal{A})$.

Proof (i) Suppose ϵ is split as $\phi\psi$, where $\psi\phi = 1_B$. Then by 3.2.1 we have

$$\phi^\circ = \psi\phi\phi^\circ \leq \psi\psi^\circ\psi\phi\phi^\circ = \psi\psi^\circ\phi^\circ = \psi(\phi\psi)^\circ = \psi\phi\psi = \psi$$

and similarly $\psi^\circ \leq \phi$, so $\psi = \phi^\circ$.

(ii) If ϵ is a core and splits as $\phi\phi^\circ$, then we have $\phi\phi^\circ \leq 1$ and $\phi^\circ\phi = 1$, so ϕ is a map; and it is split monic in $\mathcal{A}$ and hence monic in **Map**$(\mathcal{A})$.

(iii) is similar to (ii). □

If $\mathcal{A}$ is a tabular allegory, then by 3.1.4 the epimorphism g in (iii) must actually be a cover. Thus we can think of the process of splitting cores in $\mathcal{A}$ as adjoining 'missing subobjects' to **Map**$(\mathcal{A})$, and similarly splitting equivalences in $\mathcal{A}$ adjoins 'missing quotients' to **Map**$(\mathcal{A})$. We shall make this more precise shortly.

Theorem 3.3.4 *Let $\mathcal{A}$ be an allegory and $\mathcal{E}$ a class of symmetric idempotents in $\mathcal{A}$ (including all the identities). Then $\mathcal{A}[\check{\mathcal{E}}]$ has an allegory structure making the functor $I\colon \mathcal{A} \to \mathcal{A}[\check{\mathcal{E}}]$ of* 1.1.8 *a morphism of allegories. Moreover,*

(i) *If $\mathcal{A}$ has a unit, then so does $\mathcal{A}[\check{\mathcal{E}}]$ (and I preserves it).*
(ii) *If $\mathcal{A}$ is a union allegory, then so is $\mathcal{A}[\check{\mathcal{E}}]$ (and I preserves unions).*

(iii) *If $\mathcal{A}$ is a positive union allegory and $\mathcal{E}$ is either the class of all cores or the class of all equivalences in $\mathcal{A}$, then $\mathcal{A}[\check{\mathcal{E}}]$ is a positive union allegory (and I preserves finite coproducts).*

Proof The allegory structure is defined in the obvious way: the opposite of $\phi\colon \delta \looparrowright \epsilon$ is $\phi^\circ\colon \epsilon = \epsilon^\circ \looparrowright \delta^\circ = \delta$, and the intersection of ϕ and $\psi\colon \delta \looparrowright \epsilon$ is $\phi \cap \psi$. Of course, it needs verification that the latter is a morphism $\delta \looparrowright \epsilon$. But $(\phi\cap\psi)\delta \leq \phi\delta \cap \psi\delta = \phi \cap \psi$, and

$$(\phi \cap \psi)\delta = (\phi \cap \psi\delta^\circ)\delta \geq \phi\delta \cap \psi\delta = \phi \cap \psi,$$

so $(\phi \cap \psi)\delta = \phi \cap \psi$, and similarly $\epsilon(\phi \cap \psi) = \phi \cap \psi$. The validity of the modular law in $\mathcal{A}[\check{\mathcal{E}}]$ is immediate from its validity in $\mathcal{A}$. And it is immediate from the definitions that I is a morphism of allegories.

(i) If U is a unit in $\mathcal{A}$, then $1_{I(U)}$ is the greatest morphism $I(U) \looparrowright I(U)$, since I is full (and preserves the ordering on hom-posets); and if $\epsilon \in \mathcal{E}$ has domain (and codomain) A, then the composite $p_A\epsilon\colon \epsilon \looparrowright I(A) \looparrowright I(U)$ satisfies $1_\epsilon = \epsilon \leq \epsilon^\circ p_A^\circ p_A \epsilon$. Thus $I(U)$ is a unit in $\mathcal{A}[\check{\mathcal{E}}]$.

(ii) Unions in the hom-posets of $\mathcal{A}[\check{\mathcal{E}}]$, like intersections, are constructed in the 'simple-minded' way; and since composition in $\mathcal{A}$ distributes over unions, it is immediate that if ϕ and ψ are both morphisms $\delta \looparrowright \epsilon$ in $\mathcal{A}[\check{\mathcal{E}}]$, then so is $\phi \cup \psi$ (and that composition distributes over unions in $\mathcal{A}[\check{\mathcal{E}}]$).

(iii) First suppose that $\mathcal{E}$ is the class of cores. If ϵ and δ are cores on objects A and B respectively, consider the morphism $\theta = n_1\epsilon n_1^\circ \cup n_2\delta n_2^\circ\colon A + B \looparrowright A + B$ (where the n_i are the coprojections, as in 3.2.12). Clearly, we have $\theta \leq n_1 n_1^\circ \cup n_2 n_2^\circ = 1_{A+B}$, so θ is a core. Moreover,

$$\theta n_1 \epsilon = n_1 \epsilon n_1^\circ n_1 \epsilon \cup n_2 \delta n_2^\circ n_1 \epsilon = n_1 \epsilon \cup 0 = n_1 \epsilon,$$

so $n_1\epsilon$ is a morphism (in fact a map) $\epsilon \looparrowright \theta$ in $\mathcal{A}[\check{\mathcal{E}}]$. Similarly, $n_2\delta\colon \delta \looparrowright \theta$; we claim that these morphisms make θ into a coproduct of ϵ and δ in $\mathcal{A}[\check{\mathcal{E}}]$. For if $\phi\colon \epsilon \looparrowright \gamma$ and $\psi\colon \delta \looparrowright \gamma$ are two morphisms in $\mathcal{A}[\check{\mathcal{E}}]$, there is a unique $\chi\colon A + B \looparrowright C$ in $\mathcal{A}$ (where C = dom γ = cod γ) satisfying $\chi n_1 = \phi$ and $\chi n_2 = \psi$; then $\gamma\chi = \chi$ by uniqueness, and

$$\chi\theta = \phi\epsilon n_1^\circ \cup \psi\delta n_2^\circ = \phi n_1^\circ \cup \psi n_2^\circ = \chi(n_1 n_1^\circ \cup n_2 n_2^\circ) = \chi,$$

so $\chi\colon \theta \looparrowright \gamma$ in $\mathcal{A}[\check{\mathcal{E}}]$. And χ is easily seen to be the unique such morphism satisfying $\chi n_1 \epsilon = \phi$ and $\chi n_2 \delta = \psi$. Once again, it is clear from the construction that I preserves binary coproducts; and a similar argument shows that if 0 is initial in $\mathcal{A}$ then $I(0)$ is initial in $\mathcal{A}[\check{\mathcal{E}}]$.

The argument in the case when $\mathcal{E}$ is the class of equivalences is identical, except that we have to verify that θ is an equivalence rather than a core. It is

clearly reflexive and symmetric if ϵ and δ are; for transitivity, we have

$$\begin{aligned}\theta\theta &= (n_1\epsilon n_1^\circ \cup n_2\delta n_2^\circ)(n_1\epsilon n_1^\circ \cup n_2\delta n_2^\circ)\\ &= n_1\epsilon\epsilon n_1^\circ \cup 0 \cup 0 \cup n_2\delta\delta n_2^\circ\\ &\leq \theta,\end{aligned}$$

by transitivity of ϵ and δ. □

It will be observed that (iii) of the above theorem could also be proved under the hypothesis that $\mathcal{E}$ is the class of all symmetric idempotents of $\mathcal{A}$; but we shall not need this result (and cf. 3.3.12 below). For the rest of this section, we shall use the letters $\mathcal{K}$, $\mathcal{Q}$ and $\mathcal{S}$ to denote respectively the classes of cores, of equivalences and of all symmetric idempotents in a given allegory $\mathcal{A}$.

Rather remarkably, 3.3.4 is best possible in a certain sense: once we have split the symmetric idempotents, we have split all the idempotents that can be split in an allegory (which is not to say that all such idempotents are symmetric, of course).

Proposition 3.3.5 *Let ϵ be an idempotent in an allegory $\mathcal{A}$, and suppose there is a faithful morphism of allegories $F\colon \mathcal{A} \to \mathcal{B}$ such that $F(\epsilon)$ splits in $\mathcal{B}$. Then $I(\epsilon)$ splits in $\mathcal{A}[\check{\mathcal{S}}]$.*

Proof Consider the morphism $\delta = \epsilon \cap \epsilon^\circ$. δ is clearly symmetric, but it is also transitive since $\delta\delta = (\epsilon\cap\epsilon^\circ)(\epsilon\cap\epsilon^\circ) \leq \epsilon\epsilon \cap \epsilon^\circ\epsilon^\circ = \epsilon\cap\epsilon^\circ$. Hence δ is a symmetric idempotent by 3.3.2(i). We claim that $\epsilon\delta\epsilon = \epsilon$ and $\delta\epsilon\delta = \delta$; since the validity of these equations is reflected by the faithful functor F, it suffices to prove them under the assumption that ϵ is already split, say $\epsilon = \phi\psi$ with $\psi\phi = 1$. But then we have

$$\begin{aligned}\epsilon\delta\epsilon &= \phi\psi(\phi\psi \cap \psi^\circ\phi^\circ)\phi\psi\\ &\geq \phi(\psi\phi\psi \cap \phi^\circ)\phi\psi\\ &\geq \phi(\psi\phi\psi\phi \cap 1)\psi\\ &= \phi\psi = \epsilon\end{aligned}$$

since $\psi\phi = 1$, and

$$\begin{aligned}\delta\epsilon\delta &= (\phi\psi \cap \psi^\circ\phi^\circ)\phi\psi(\phi\psi \cap \psi^\circ\phi^\circ)\\ &\leq (\phi \cap \psi^\circ\phi^\circ\psi^\circ)\psi\phi\psi\phi(\psi \cap \phi^\circ\psi^\circ\phi^\circ)\\ &= (\phi\cap\psi^\circ)(\psi\cap\phi^\circ)\\ &\leq \phi\psi \cap \psi^\circ\phi^\circ = \delta;\end{aligned}$$

the reverse inequalities are trivial since $\epsilon = \epsilon^3 \geq \epsilon\delta\epsilon$ and $\delta = \delta^3 \leq \delta\epsilon\delta$. It now follows that any morphism of allegories which splits δ will also split ϵ; for if $\delta = \alpha\beta$ with $\beta\alpha = 1$, then $\epsilon = (\epsilon\alpha)(\beta\epsilon)$ and

$$(\beta\epsilon)(\epsilon\alpha) = \beta\epsilon\alpha = (\beta\delta)\epsilon(\delta\alpha) = \beta\delta\alpha = \beta\alpha = 1\,.$$

□

It follows easily from 3.3.3(ii) that a core has a splitting ff° iff it has a tabulation (f, f); so, in a tabular allegory, all cores must split. However, this condition is not sufficient for tabularity: consider the allegory of all sets of cardinality at most 10 and relations between them. In this, all cores (indeed, all symmetric idempotents) split; but not every morphism has a tabulation, essentially because the category of maps lacks products.

We say an allegory is *pre-tabular* if every morphism is contained in one which has a tabulation. If the hom-posets of $\mathcal{A}$ have top elements (e.g. if $\mathcal{A}$ has a unit, cf. 3.2.9), then this is equivalent to demanding that these top elements have tabulations. Further, if the greatest morphism $A \mathrel{\mapstochar\rightarrow} B$ has a tabulation, the latter is a product of A and B in $\mathbf{Map}(\mathcal{A})$, by 3.2.4. (The converse may fail: if A and B have a product in $\mathbf{Map}(\mathcal{A})$, the most we can say about the morphism $A \mathrel{\mapstochar\rightarrow} B$ tabulated by the product projections is that it is the greatest which can be factored in the form gf°.)

Proposition 3.3.6

(i) *An allegory is tabular iff it is pre-tabular and cores split.*

(ii) *If $\mathcal{A}$ is a pre-tabular allegory, then $\mathcal{A}[\check{\mathcal{K}}]$ is tabular, and the functor $I: \mathcal{A} \to \mathcal{A}[\check{\mathcal{K}}]$ is universal among morphisms from $\mathcal{A}$ to tabular allegories.*

Proof (i) One direction is obvious. Conversely, suppose $\mathcal{A}$ is pre-tabular and cores split; let $\phi: A \mathrel{\mapstochar\rightarrow} B$ be a morphism and (f, g) a tabulation such that $\phi \leq gf^\circ$. Then, if $T = \mathrm{dom}\ f = \mathrm{dom}\ g$, the domain of the maps in a tabulation of ϕ 'ought to be' a subobject of T, so we should be able to obtain it by splitting a suitable core on T. Consider the core $1_T \cap g^\circ \phi f$; suppose it splits as hh°. We claim that (fh, gh) tabulates ϕ. First, we have

$$(gh)(fh)^\circ = ghh^\circ f^\circ \leq gg^\circ \phi f f^\circ \leq \phi$$

and

$$\phi = \phi \cap gf^\circ \leq g(g^\circ \phi \cap f^\circ) \leq g(g^\circ \phi f \cap 1_T)f^\circ = ghh^\circ f^\circ\ .$$

Then

$$\begin{aligned}(fh)^\circ(fh) \cap (gh)^\circ(gh) &= h^\circ f^\circ fh \cap h^\circ g^\circ gh \\ &= h^\circ(f^\circ f \cap g^\circ g)h \\ &= h^\circ h = 1,\end{aligned}$$

using 3.1.6 and its dual at the second step.

(ii) Let $\phi: \epsilon \mathrel{\mapstochar\rightarrow} \delta$ be a morphism in $\mathcal{A}[\check{\mathcal{K}}]$, and let (f, g) be a tabulation of a morphism which contains ϕ in $\mathcal{A}$. Let $\theta = f^\circ \epsilon f \cap g^\circ \delta g$; then $\theta \leq f^\circ f \cap g^\circ g = 1$, so θ is a core and we have $\epsilon f \theta: \theta \mathrel{\mapstochar\rightarrow} \epsilon$, $\delta g \theta: \theta \mathrel{\mapstochar\rightarrow} \delta$ in $\mathcal{A}[\check{\mathcal{K}}]$. Moreover,

$$\epsilon f \theta \theta^\circ f^\circ \epsilon^\circ = \epsilon f \theta f^\circ \epsilon \leq \epsilon f f^\circ \epsilon \leq \epsilon$$

and

$$\theta^\circ f^\circ \epsilon^\circ \epsilon f \theta = \theta f \epsilon f^\circ \theta \geq \theta\theta\theta = \theta,$$

so $\epsilon f\theta$ is a map in $\mathcal{A}[\check{\mathcal{K}}]$, and similarly $\delta g\theta$ is a map. Now

$$\begin{aligned}(\delta g\theta)(\epsilon f\theta)^\circ &= \delta g\theta f^\circ\epsilon \\ &= \delta g(f^\circ\epsilon f \cap g^\circ\delta g)f^\circ\epsilon \\ &\geq (\delta g f^\circ\epsilon f \cap g)f^\circ\epsilon \\ &\geq \delta g f^\circ \cap g f^\circ\epsilon \\ &\geq \delta\phi \cap \phi\epsilon = \phi \cap \phi = \phi\end{aligned}$$

and

$$\begin{aligned}(\epsilon f\theta)^\circ(\epsilon f\theta) \cap (\delta g\theta)^\circ(\delta g\theta) &= \theta f^\circ\epsilon f\theta \cap \theta g^\circ\delta g\theta \\ &= \theta(f^\circ\epsilon f \cap g^\circ\delta g)\theta \\ &\leq \theta(f^\circ f \cap g^\circ g)\theta = \theta\end{aligned}$$

(where we have again used 3.1.6 at the second step), so $(\epsilon f\theta, \delta g\theta)$ tabulates a morphism containing ϕ. Thus $\mathcal{A}[\check{\mathcal{K}}]$ is pre-tabular. But any core in $\mathcal{A}[\check{\mathcal{K}}]$ is also a core in $\mathcal{A}$ and so splits in $\mathcal{A}[\check{\mathcal{K}}]$; thus $\mathcal{A}[\check{\mathcal{K}}]$ is tabular by part (i). The last assertion is also immediate from (i), since if $T\colon \mathcal{A} \to \mathcal{B}$ is any morphism from $\mathcal{A}$ to a tabular allegory, it must send cores in $\mathcal{A}$ to (split) cores in $\mathcal{B}$. □

Example 3.3.7 The allegory $\mathbf{Mat}_f(L)$ of 3.2.2(*b*) is pre-tabular; the greatest morphism $A \looparrowright B$ is the matrix with all entries 1, which lies in the sub-allegory $\mathbf{Mat}_f(\mathbf{2})$ and has a tabulation there. So if we split its cores we obtain a tabular allegory, which has a unit and is a positive union allegory since $\mathbf{Mat}_f(L)$ enjoys these properties. What is the corresponding (positive coherent) category of maps? Working through the definitions, we see that a core ϵ in $\mathbf{Mat}_f(L)$ is a diagonal matrix, i.e. one satisfying $\epsilon(a, a') = 0$ whenever $a \neq a'$, so an object of $\mathbf{Mat}_f(L)[\check{\mathcal{K}}]$ may be identified with a pair (A, α) where A is a finite set and $\alpha\colon A \to L$ an arbitrary function. A matrix $\phi\colon B \times A \to L$ defines a morphism $(A, \alpha) \looparrowright (B, \beta)$ iff $\phi(b, a) \leq \beta(b) \wedge \alpha(a)$ for all a and b; it is a map if, in addition, its ath column is a disjoint decomposition of $\alpha(a)$ for each a, i.e. $\phi(b, a) \wedge \phi(b', a) = 0$ whenever $b \neq b'$, and

$$\bigvee\{\phi(b, a) \mid b \in B\} = \alpha(a) .$$

Thus we see that $\mathbf{Map}(\mathbf{Mat}_f(L)[\check{\mathcal{K}}])$ is none other than (a category equivalent to) the positivization $\mathbf{Pos}(L)$ of L, which we considered in 1.4.6; equivalently, $\mathbf{Mat}_f(L)[\check{\mathcal{K}}]$ is equivalent to $\mathbf{Rel}(\mathbf{Pos}(L))$.

Example 3.3.8 The core-splitting construction may also be used to give an alternative description of the regularization of a cartesian category, constructed in 1.3.9. We recall from 3.1.1 that if $\mathcal{C}$ is merely cartesian, then relations in $\mathcal{C}$ do not in general form a category, since they do not possess an associative composition. However, we do always have a bicategory (cf. Section B1.1) $\mathfrak{Span}(\mathcal{C})$ whose objects are those of $\mathcal{C}$, whose morphisms $A \to B$ are *spans* in $\mathcal{C}$ (that is, diagrams of the form $(A \leftarrow T \to B)$, without any requirement that the pair be jointly monic), and whose 2-cells $(A \leftarrow T \to B) \to (A \leftarrow T' \to B)$ are morphisms

$T \to T'$ commuting with the given morphisms to A and B. (Equivalently, the hom-category $\mathfrak{Span}(\mathcal{C})\,(A,B)$ may be regarded as the slice category $\mathcal{C}/A \times B$.) The composite of two spans $(A \leftarrow T \to B)$ and $(B \leftarrow U \to C)$ is defined by forming the pullback

$$\begin{tikzcd} V \arrow[r] \arrow[d] & U \arrow[d] \\ T \arrow[r] & B \end{tikzcd}$$

and composing its projections with $T \to A$ and $U \to C$. (Note that this composition is *not* strictly associative in general; thus $\mathfrak{Span}(\mathcal{C})$ is only a bicategory, not a 2-category.) If we form the preorder reflections of the hom-categories $\mathfrak{Span}(\mathcal{C})\,(A,B)$ (that is, we form their quotients by the equivalence relations which identify all parallel pairs of 2-cells), we still do not have a 2-category, but if we further identify all isomorphic pairs of 1-cells, then composition does become strictly associative. Thus we have a locally ordered 2-category $\overline{\mathfrak{Span}}(\mathcal{C})$ whose objects are those of $\mathcal{C}$ and whose morphisms are equivalence classes of spans (under the equivalence relation which identifies two spans if there exist morphisms between them in both directions), with ordering defined by the relation that there exists a morphism of spans from (a span representing) the first morphism to the second.

Moreover, $\overline{\mathfrak{Span}}(\mathcal{C})$ is an allegory: binary intersections of morphisms are given by (the equivalence classes of) pullbacks in $\mathcal{C}/A \times B$, the opposite of a morphism is defined just as in $\mathbf{Rel}(\mathcal{C})$, and the modular law is proved in the same way that we proved it for relations in 3.1.5 (though the proof is easier, because we do not have to take images). Also, the terminal object of $\mathcal{C}$ is easily seen to be a unit in $\overline{\mathfrak{Span}}(\mathcal{C})$; and, for any morphism $f\colon A \to B$ of $\mathcal{C}$, the morphism $f_\bullet\colon A \nrightarrow B$ of $\overline{\mathfrak{Span}}(\mathcal{C})$ represented by the graph of f (that is, the span $(1,f)\colon A \to A \times B$) is a map (though we do not claim that all maps in $\overline{\mathfrak{Span}}(\mathcal{C})$ are of this form), from which it follows easily that the top element of $\overline{\mathfrak{Span}}(\mathcal{C})\,(A,B)$ (which is of course represented by the span consisting of the product projections from $A \times B$) has a tabulation. So $\overline{\mathfrak{Span}}(\mathcal{C})$ is pre-tabular; hence if we split its cores we obtain a tabular allegory with a unit – equivalently, by 3.2.10, we obtain an allegory equivalent to the category of relations of a regular category.

Now if $F\colon \mathcal{C} \to \mathcal{D}$ is any cartesian functor from $\mathcal{C}$ to a regular category, we obtain a morphism of allegories from $\overline{\mathfrak{Span}}(\mathcal{C})$ to $\mathbf{Rel}(\mathcal{D})$ by sending the equivalence class of a span $(T \to A \times B)$ to the image in $\mathcal{D}$ of $(FT \to FA \times FB)$. Thus the universal property of 3.3.6(ii) tells us that $\mathbf{Map}(\overline{\mathfrak{Span}}(\mathcal{C})[\check{\mathcal{K}}])$ is universal amongst regular categories into which $\mathcal{C}$ can be mapped by cartesian functors; hence it must be equivalent to $\mathbf{Reg}(\mathcal{C})$ as constructed in 1.3.9. (In fact it is not hard to show directly that $\overline{\mathfrak{Span}}(\mathcal{C})[\check{\mathcal{K}}] \simeq \mathbf{Rel}(\mathbf{Reg}(\mathcal{C}))$.)

Example 3.3.7 shows that $\mathbf{Map}(\mathcal{A}[\check{\mathcal{K}}])$ is not in general effective as a regular category; to achieve this, we need to split some more symmetric idempotents, namely the equivalences. We call an allegory *effective* if its equivalences split.

Proposition 3.3.9

(i) *A tabular allegory with a unit is effective iff its category of maps is effective regular.*

(ii) *For any allegory $\mathcal{A}$, $\mathcal{A}[\check{\mathcal{Q}}]$ is effective, and the functor $I\colon \mathcal{A} \to \mathcal{A}[\check{\mathcal{Q}}]$ is universal among morphisms from $\mathcal{A}$ to effective allegories. Moreover, if $\mathcal{A}$ is tabular, so is $\mathcal{A}[\check{\mathcal{Q}}]$.*

Proof (i) We have already observed that equivalences in $\mathbf{Rel}(\mathcal{C})$, for a regular category $\mathcal{C}$, are the same thing as equivalence relations in $\mathcal{C}$, as defined in 1.3.6. Moreover, an equivalence relation $(a,b)\colon R \rightrightarrows A$ is the kernel-pair of a cover $q\colon A \twoheadrightarrow B$ iff it is a tabulation of $q^\bullet q_\bullet\colon A \looparrowright A$ (by the proof of 3.2.7), iff it has a splitting in $\mathbf{Rel}(\mathcal{C})$ (by 3.3.3(iii)).

(ii) As in 3.3.6(ii), it is clear that any equivalence in $\mathcal{A}[\check{\mathcal{Q}}]$ is an equivalence in $\mathcal{A}$ and hence splits in $\mathcal{A}[\check{\mathcal{Q}}]$; so the first two assertions are immediate. Now suppose $\mathcal{A}$ is tabular; let $\phi\colon \epsilon \looparrowright \delta$ be a morphism of $\mathcal{A}[\check{\mathcal{Q}}]$, and let (f,g) be a tabulation of ϕ in $\mathcal{A}$. Define $\theta = f^\circ\epsilon f \cap g^\circ\delta g$; then $\theta \geq f^\circ f \cap g^\circ g = 1$, so θ is reflexive. It is also clearly symmetric, and

$$\begin{aligned}\theta\theta &= (f^\circ\epsilon f \cap g^\circ\delta g)(f^\circ\epsilon f \cap g^\circ\delta g)\\ &\leq f^\circ\epsilon f f^\circ\epsilon f \cap g^\circ\delta g g^\circ\delta g\\ &\leq f^\circ\epsilon f \cap g^\circ\delta g = \theta\end{aligned}$$

since f and g are maps, so θ is transitive. Now $\epsilon f\theta \leq \epsilon f f^\circ\epsilon f \leq \epsilon f$, but we also have $\epsilon f\theta \geq \epsilon f$ since θ is reflexive; so ϵf is a morphism $\theta \looparrowright \epsilon$ in $\mathcal{A}[\check{\mathcal{Q}}]$. In fact it is easily seen to be a map, and similarly $\delta g\colon \theta \to \delta$. Now $\phi = \delta\phi\epsilon = (\delta g)(\epsilon f)^\circ$, and

$$(\epsilon f)^\circ(\epsilon f) \cap (\delta g)^\circ(\delta g) = f^\circ\epsilon f \cap g^\circ\delta g = \theta,$$

so $(\epsilon f, \delta g)$ is a tabulation of $\phi\colon \epsilon \looparrowright \delta$ in $\mathcal{A}[\check{\mathcal{Q}}]$. □

Corollary 3.3.10 *Let $\mathcal{C}$ be a regular category. Then there exists an effective regular category $\mathbf{Eff}(\mathcal{C})$ and a full and faithful functor $\mathcal{C} \to \mathbf{Eff}(\mathcal{C})$ which is universal among regular functors from $\mathcal{C}$ to effective regular categories. Moreover, if $\mathcal{C}$ is a (positive) coherent category, so is $\mathbf{Eff}(\mathcal{C})$, and the functor $\mathcal{C} \to \mathbf{Eff}(\mathcal{C})$ is then coherent.*

Proof Define $\mathbf{Eff}(\mathcal{C})$ to be $\mathbf{Map}(\mathbf{Rel}(\mathcal{C})[\check{\mathcal{Q}}])$, and translate 3.3.9 via the equivalence of 3.2.10. (Note, incidentally, that a full and faithful morphism of allegories automatically restricts to a full and faithful functor between categories of maps.) □

Remark 3.3.11 Let $\mathcal{C}$ be a regular category having coequalizers of equivalence relations. Then the inclusion $\mathcal{C} \to \mathbf{Eff}(\mathcal{C})$ has a left adjoint. For if $\phi \colon A \looparrowright A$ is an equivalence in $\mathbf{Rel}(\mathcal{C})$ – that is, an object of $\mathbf{Eff}(\mathcal{C})$ – and $A \leftarrow R \to A$ is a tabulation of ϕ, then it is easy to see that maps $\phi \to \iota_B$ in $\mathbf{Rel}(\mathcal{C})[\check{\mathcal{Q}}]$ are exactly the maps $A \to B$ in $\mathcal{C}$ having equal composites with $R \rightrightarrows A$. So we obtain the left adjoint by sending ϕ to the coequalizer of this pair in $\mathcal{C}$.

We note that the two splitting constructions, of cores and equivalences, could have been combined into one:

Lemma 3.3.12 *In an allegory, all symmetric idempotents split iff cores and equivalences split.*

Proof One direction is trivial. Conversely, suppose that cores and equivalences split, and let $\epsilon \colon A \looparrowright A$ be an arbitrary symmetric idempotent. Then $\epsilon \cap 1_A$ is a core, and so we can split it as ff° for some $f \colon B \to A$. Now consider $\delta = f^\circ \epsilon f \colon B \looparrowright B$. We have

$$\delta \geq f^\circ(\epsilon \cap 1_A)f = f^\circ f f^\circ f = 1_B,$$

so δ is reflexive; it is clearly symmetric, and it is transitive since

$$\delta\delta = f^\circ \epsilon f f^\circ \epsilon f \leq f^\circ \epsilon\epsilon\epsilon f = f^\circ \epsilon f .$$

Thus δ is an equivalence, and we can split it as $g^\circ g$ for some $g \colon B \to C$. Now $fg^\circ g f^\circ = f\delta f^\circ = ff^\circ \epsilon f f^\circ \leq \epsilon$ since $ff^\circ \leq 1_A$, but we also have

$$\begin{aligned} ff^\circ \epsilon f f^\circ &= (\epsilon \cap 1_A)\epsilon(\epsilon \cap 1_A) \\ &= (\epsilon\epsilon^\circ \cap 1_A)\epsilon\epsilon(\epsilon^\circ\epsilon \cap 1_A) \\ &\geq (\epsilon \cap \epsilon)(\epsilon \cap \epsilon) = \epsilon; \end{aligned}$$

and $gf^\circ f g^\circ = gg^\circ = 1_C$, so (gf°, fg°) is a splitting of ϵ. □

Thus, for any allegory $\mathcal{A}$, we may identify $\mathcal{A}[\check{\mathcal{K}}][\check{\mathcal{Q}}]$ with $\mathcal{A}[\check{\mathcal{S}}]$; though one should beware that in this identification $\mathcal{Q}$ denotes the class of all equivalences in $\mathcal{A}[\check{\mathcal{K}}]$, and not just those in $\mathcal{A}$. In particular, we obtain

Corollary 3.3.13

(i) *For any cartesian category $\mathcal{C}$, $\mathbf{Map}(\overline{\mathfrak{Span}}(\mathcal{C})[\check{\mathcal{S}}])$ is an effective regular category, and there is a full embedding $\mathcal{C} \to \mathbf{Map}(\overline{\mathfrak{Span}}(\mathcal{C})[\check{\mathcal{S}}])$ which is universal amongst cartesian functors from $\mathcal{C}$ to effective regular categories.*

(ii) *For any distributive lattice L, the category $\mathbf{Map}(\mathbf{Mat}_f(L)[\check{\mathcal{S}}])$ is a pretopos, and there is an embedding $L \to \mathbf{Map}(\mathbf{Mat}_f(L)[\check{\mathcal{S}}])$ which is universal amongst coherent functors from L to pretoposes.*

Proof The first assertion is immediate from 3.3.8 and 3.3.10; the second from 3.3.7 and 3.3.10. □

We may describe the pretopos of 3.3.13(ii) explicitly as follows. A symmetric idempotent in $\mathbf{Mat}_f(L)$ is a function $\epsilon\colon A \times A \to L$ satisfying $\epsilon(a_1, a_2) = \epsilon(a_2, a_1)$ and $\epsilon(a_1, a_2) \wedge \epsilon(a_2, a_3) \leq \epsilon(a_1, a_3)$ for all $a_1, a_2, a_3 \in A$; that is, it is an 'L-valued partial equivalence relation' on A. ('Partial' because we don't demand the reflexivity condition $\epsilon(a, a) = 1$.) And a map $(A, \epsilon) \to (B, \delta)$ is a function $\phi\colon B \times A \to L$ satisfying

$$\begin{gathered}\phi(b, a) \leq \delta(b, b) \wedge \epsilon(a, a),\\ \phi(b_1, a_1) \wedge \delta(b_1, b_2) \wedge \epsilon(a_1, a_2) \leq \phi(b_2, a_2),\\ \phi(b_1, a) \wedge \phi(b_2, a) \leq \delta(b_1, b_2),\end{gathered}$$

and

$$\epsilon(a, a) \leq \bigvee \{\phi(b, a) \mid b \in B\}$$

for all $a, a_1, a_2 \in A$ and all $b, b_1, b_2 \in B$. As usual, composition of maps is defined by matrix muliplication; it is not hard to verify directly that the four conditions above are stable under composition.

In the case when L is a frame, and with the finiteness restrictions on A and B removed, the above formulae define the category of maps of $\mathbf{Mat}(L)[\check{S}]$; once again this category is a pretopos, and we shall see in the next section that it is actually a topos. We denote it by $\mathbf{Set}(L)$, and call it the category of *L-valued sets*.

Suggestion for further reading: Freyd & Scedrov [381].

A3.4 Division allegories and power allegories

Next, we consider the extra structure on an allegory which corresponds to that of a Heyting category.

Definition 3.4.1 An allegory $\mathcal{A}$ is called a *division allegory* if, for each $\phi\colon A \looparrowright B$ and each object C, the order-preserving map $(-)\phi\colon \mathcal{A}(B, C) \to \mathcal{A}(A, C)$ has a right adjoint, which we call *right division* by ϕ and denote $(-)/\phi$.

Of course, the anti-involution ensures that if we have right division we also have *left division* $\phi\backslash(-)$ (right adjoint to $\phi(-)$), since $\phi\backslash\psi$ may be defined as $(\psi^\circ/\phi^\circ)^\circ$. We note that the operation $(-)/\phi$ satisfies the identities

$$\begin{gathered}(\psi \cap \chi)/\phi = \psi/\phi \cap \chi/\phi,\\ \psi\phi/\phi \geq \psi\end{gathered}$$

and

$$(\chi/\phi)\phi \leq \chi$$

whenever the expressions involved in them are defined; and indeed these identities suffice to characterize $(-)/\phi$. A further useful identity is the associativity of double division: $(\phi\backslash\psi)/\chi = \phi\backslash(\psi/\chi)$ whenever this makes sense, for each of the

two expressions defines the largest θ satisfying $\phi\theta\chi \leq \psi$. Similarly, we may verify that $(\phi/\psi)/\chi = \phi/(\chi\psi)$ whenever this makes sense, and that $\phi(\psi/\chi) \leq (\phi\psi)/\chi$. Also, if the division operations exist, then composition automatically distributes over any unions which may exist in the hom-posets of $\mathcal{A}$; conversely, a geometric allegory (as defined at the end of Section A3.2) is a division allegory, by the poset version of the adjoint functor theorem.

Lemma 3.4.2 *Let $\mathcal{A}$ be a tabular allegory with a unit, whose hom-posets have finite unions. Then $\mathcal{A}$ is a division allegory iff* **Map**$(\mathcal{A})$ *is a Heyting category.*

Proof As usual, this is entirely straightforward if we translate it, via 3.2.10, into a statement about a regular category $\mathcal{C}$ and the corresponding allegory **Rel**$(\mathcal{C})$. If (f,g) is a tabulation of $\phi\colon A \rightarrow B$, then $(-)\phi$ may be regarded as the composite

$$\mathrm{Sub}(B \times C) \xrightarrow{(g\times 1)^*} \mathrm{Sub}(T \times C) \xrightarrow{\exists_{(f\times 1)}} \mathrm{Sub}(A \times C),$$

and so has a right adjoint $\forall_{(g\times 1)} \circ (f \times 1)^*$ if $\mathcal{C}$ is a Heyting category. The converse is obtained by identifying $\mathrm{Sub}_{\mathcal{C}}(A)$ with **Rel**$(\mathcal{C})(A,1)$. □

Lemma 3.4.3 *Let $\mathcal{A}$ be a division allegory, $\mathcal{E}$ a class of symmetric idempotents of $\mathcal{A}$. Then $\mathcal{A}[\check{\mathcal{E}}]$ is a division allegory, and the canonical functor $I\colon \mathcal{A} \to \mathcal{A}[\check{\mathcal{E}}]$ preserves division.*

Proof Let $\phi\colon \epsilon \rightarrow \delta$ and $\psi\colon \epsilon \rightarrow \gamma$ be morphisms of $\mathcal{A}[\check{\mathcal{E}}]$. Then $\gamma(\psi/\phi)\delta$ is certainly a morphism $\delta \rightarrow \gamma$; and

$$\gamma(\psi/\phi)\delta\phi = \gamma(\psi/\phi)\phi \leq \gamma\psi = \psi\,.$$

But if $\chi\colon \delta \rightarrow \gamma$ is any morphism satisfying $\chi\phi \leq \psi$, then we have $\chi \leq \psi/\phi$ in $\mathcal{A}$, whence $\chi = \gamma\chi\delta \leq \gamma(\psi/\phi)\delta$. So $\gamma(-/\phi)\delta$ is right adjoint to

$$\mathcal{A}[\check{\mathcal{E}}](\delta,\gamma) \xrightarrow{(-)\phi} \mathcal{A}[\check{\mathcal{E}}](\epsilon,\gamma)\,.$$

The second assertion is immediate. □

If L is a Heyting algebra, then **Mat**$_f(L)$ is a division allegory: if $\phi\colon A \rightarrow B$, $\psi\colon A \rightarrow C$ and $\chi\colon B \rightarrow C$ are three L-valued matrices, we have $\chi\phi \leq \psi$ iff $\chi(c,b) \wedge \phi(b,a) \leq \psi(c,a)$ for all a,b,c, iff $\chi(c,b) \leq (\phi(b,a) \Rightarrow \psi(c,a))$ for all a,b,c, and so we may define the matrix ψ/ϕ by

$$\psi/\phi(c,b) = \bigwedge\{(\phi(b,a) \Rightarrow \psi(c,a)) \mid a \in A\}\,.$$

Thus we may deduce from the two preceding results that **Pos**(L) and **Eff**(**Pos**(L)) are Heyting categories. Similarly, if L is a complete Heyting algebra

(equivalently, a frame), then the category $\mathbf{Set}(L)$ defined at the end of the last section is a Heyting category.

In fact, as we remarked earlier, $\mathbf{Set}(L)$ is a topos; so the next step is to seek a condition on an allegory which is equivalent, in the tabular case, to its category of maps being a topos. In one sense, there is a simple answer:

Lemma 3.4.4 *Let $\mathcal{C}$ be a regular category. Then $\mathcal{C}$ is a topos iff the functor $\mathcal{C} \to \mathbf{Rel}(\mathcal{C})$ which sends f to $f_\bullet$ has a right adjoint.*

Proof Such a right adjoint, if it exists, will send each object A of $\mathcal{C}$ to an object PA equipped with a relation $\in_A$: $PA \looparrowright A$ such that any relation $\phi\colon B \looparrowright A$ can be factored as $\in_A f_\bullet$ for a unique map $f\colon B \to PA$. But, if we identify relations with subobjects of products, $\in_A f_\bullet$ is just another name for $(f \times 1_A)^*(\in_A)$; so this is exactly the definition of a power object as given in Section A2.1. □

Thus we could define a power allegory as an allegory $\mathcal{A}$ for which the inclusion $\mathbf{Map}(\mathcal{A}) \to \mathcal{A}$ has a right adjoint. It turns out, however, that in the presence of the division operation this description may be further simplified, since the 'naming' operation from morphisms to maps is definable in terms of the other structure. (And since a topos is a Heyting category by 2.3.5, it is not unreasonable to demand that the division operation should be present in a power allegory.) Before giving the definition, however, we need to introduce a 'symmetrized' version of the division operation: if $\phi\colon B \looparrowright A$ and $\psi\colon C \looparrowright A$ are morphisms in a division allegory with the same codomain, we write $(\phi|\psi)$ for

$$(\phi\backslash\psi) \cap (\phi^\circ/\psi^\circ) = (\phi\backslash\psi) \cap (\psi\backslash\phi)^\circ \colon C \looparrowright B\ .$$

(In the allegory $\mathbf{Rel(Set)}$, $(\phi|\psi)$ relates a pair of elements (c, b) iff they are related by ψ and ϕ respectively to exactly the same elements of A.) In general, we note that $(\psi|\phi) = (\phi|\psi)^\circ$; and if $\chi\colon D \looparrowright A$ is a third morphism, then

$$\begin{aligned}(\phi|\psi)(\psi|\chi) &\leq (\phi\backslash\psi)(\psi\backslash\chi) \cap (\phi^\circ/\psi^\circ)(\psi^\circ/\chi^\circ)\\ &\leq (\phi\backslash\chi) \cap (\phi^\circ/\chi^\circ) = (\phi|\chi)\ .\end{aligned}$$

Also, since $\phi\backslash\phi \geq 1$ for any ϕ, it is easy to see that $(\phi|\phi)$ is always an equivalence. Conversely, if ϵ is an equivalence in a division allegory, then transitivity yields $\epsilon \leq \epsilon\backslash\epsilon$, whence $\epsilon \leq (\epsilon|\epsilon)$ by symmetry; but we also have $\epsilon = 1\backslash\epsilon \geq \epsilon\backslash\epsilon \geq (\epsilon|\epsilon)$, since ϵ is reflexive and $-\backslash-$ is order-reversing in its first variable. Thus the equivalences in a division allegory are exactly the morphisms of the form $(\phi|\phi)$.

Definition 3.4.5 A division allegory $\mathcal{A}$ is called a *power allegory* if there is an operation assigning to each object A a morphism $\in_A$: $PA \looparrowright A$ satisfying $(\in_A | \in_A) = 1_{PA}$ and

$$1_B \leq (\phi\backslash \in_A)(\in_A \backslash\phi)$$

for any $\phi\colon B \looparrowright A$.

If $\mathcal{A}$ is a power allegory, we shall call PA the *power object* of A, and write $\ulcorner\phi\urcorner$ for $(\in_A|\phi)\colon B \nrightarrow PA$, which we call the *name* of the morphism $\phi\colon B \nrightarrow A$. (Intuitively, we can think of the two conditions in the definition as allegorical versions of the extensionality and comprehension axioms: in **Rel**(**Set**), the condition $(\in_A \mid \in_A) = 1_{PA}$ says that two subsets of A which have the same members are equal, and the second condition says that for any $\phi\colon B \nrightarrow A$ and any $b \in B$ there is a subset of A whose members are precisely the elements of A related to b by ϕ.) To justify the terminology just introduced, we need

Lemma 3.4.6 (i) *For any ϕ, $\ulcorner\phi\urcorner$ is a map.*

ii) *The assignment $A \mapsto PA$ extends to a functor $\mathcal{A} \to$ **Map**$(\mathcal{A})$, right adjoint to the inclusion.*

Proof (i) We have $\ulcorner\phi\urcorner\ulcorner\phi\urcorner^\circ = (\in_A|\phi)(\phi|\in_A) \leq (\in_A|\in_A) = 1_{PA}$, and

$$\begin{aligned}\ulcorner\phi\urcorner^\circ\ulcorner\phi\urcorner &= ((\in_A\backslash\phi)^\circ \cap (\phi\backslash\in_A))((\in_A\backslash\phi)\cap(\phi\backslash\in_A)^\circ\cap(\in_A\backslash\phi))\\ &\geq 1_B \cap ((\in_A\backslash\phi)^\circ \cap (\phi\backslash\in_A))(\in_A\backslash\phi)\\ &\geq 1_B \cap 1_B \cap (\phi\backslash\in_A)(\in_A\backslash\phi) = 1_B\end{aligned}$$

by two applications of the modular law.

(ii) First, we claim that $\in_A\ulcorner\phi\urcorner = \phi$ for any $\phi\colon B \nrightarrow A$. For

$$\in_A\ulcorner\phi\urcorner \leq \in_A(\in_A\backslash\phi) \leq \phi,$$

and

$$\phi \leq \phi\ulcorner\phi\urcorner^\circ\ulcorner\phi\urcorner \leq \phi(\phi\backslash\in_A)\ulcorner\phi\urcorner \leq \in_A\ulcorner\phi\urcorner .$$

Now if $f\colon B \to PA$ is any map satisfying $\in_A f = \phi$, we have $f \leq (\in_A\backslash\phi)$, and also $f\phi^\circ = ff^\circ \in_A^\circ \leq \in_A^\circ$, whence $f \leq (\in_A^\circ/\phi^\circ)$. So $f \leq \ulcorner\phi\urcorner$, whence $f = \ulcorner\phi\urcorner$ by 3.2.3(ii). Thus $f \mapsto \in_A f$ is a bijection

$$\mathbf{Map}(\mathcal{A})(B, PA) \longrightarrow \mathcal{A}(B, A);$$

the remaining details are straightforward. □

Corollary 3.4.7 *If $\mathcal{A}$ is a tabular allegory with a unit, then $\mathcal{A}$ is a power allegory iff* **Map**$(\mathcal{A})$ *is a topos.*

Proof One direction is immediate from 3.4.4 and 3.4.6. Conversely, if $\mathcal{C}$ is a topos, then **Rel**$(\mathcal{C})$ is a division allegory by 2.3.5 and 3.4.2, and it is not hard to verify directly that the membership relation on A, as defined in Section A2.1, satisfies the identities of 3.4.5. □

Not every power allegory is tabular: for a simple counterexample, take the full subcategory of **Rel**(**Set**$_f$) whose objects are sets whose cardinality is a power of 2. (This example is pre-tabular, since the corresponding category of maps has finite products, but it can easily be modified to make it not even pre-tabular.)

However, once we split the cores in this example, we obtain an allegory which is not only tabular but effective, and this is typical:

Proposition 3.4.8 *Let $\mathcal{A}$ be a power allegory. Then*

(i) *If $\mathcal{K}$ is the class of cores in $\mathcal{A}$, $\mathcal{A}[\check{\mathcal{K}}]$ is a power allegory.*

(ii) *Every equivalence in $\mathcal{A}$ can be factored as $f^\circ f$ where f is a map.*

(iii) *If cores split in $\mathcal{A}$, so do equivalences (and hence $\mathcal{A}[\check{\mathcal{K}}] \simeq \mathcal{A}[\check{\mathcal{S}}]$, where $\mathcal{S}$ is the class of all symmetric idempotents in $\mathcal{A}$).*

(iv) *If $\mathcal{A}$ is nonempty and effective and its hom-posets have bottom elements, then it has a unit.*

Proof (i) Let $\delta\colon A \looparrowright A$ be a core in $\mathcal{A}$. Define $\pi = \ulcorner\delta\in_A\urcorner \cap 1_{PA}$; then π is a core on PA, and we shall show that it serves as a power object for δ in $\mathcal{A}[\check{\mathcal{K}}]$. First, we have $\in_A\pi \leq \in_A\ulcorner\delta\in_A\urcorner = \delta\in_A$, and so

$$\in_A\pi = \in_A\pi\pi \leq \delta\in_A\pi \leq \in_A\pi$$

using the fact that both δ and π are cores. Thus $\in_A\pi\colon \pi \looparrowright \delta$ in $\mathcal{A}[\check{\mathcal{K}}]$; we shall show that it satisfies the conditions of 3.4.5. First,

$$\in_A\pi(\in_A\pi\backslash\in_A\pi)\pi \leq \in_A\pi\pi \leq \delta\in_A,$$

so $\pi(\in_A\pi\backslash\in_A\pi)\pi \leq (\in_A\backslash\delta\in_A)$. Thus

$$\begin{aligned}\pi(\in_A\pi\backslash\in_A\pi)\pi \cap (\pi(\in_A\pi\backslash\in_A\pi)\pi)^\circ &\leq (\in_A\backslash\delta\in_A)\cap(\in_A\backslash\delta\in_A)^\circ\\ &\leq (\in_A\backslash\delta\in_A)\cap(\in_A\backslash\in_A)^\circ\\ &= (\in_A\backslash\delta\in_A)\cap(\delta\in_A\backslash\in_A)^\circ\cap(\in_A\backslash\in_A)\\ &\quad\ \cap(\in_A\backslash\in_A)^\circ\\ &= \ulcorner\delta\in_A\urcorner\cap 1_{PA} = \pi,\end{aligned}$$

using the fact that δ is a core and the order-preserving properties of $-\backslash-$.

Now if $\phi\colon \gamma \looparrowright \delta$ is any morphism of $\mathcal{A}[\check{\mathcal{K}}]$, we have $\delta\phi = \phi$, and hence

$$\in_A\ulcorner\delta\in_A\urcorner\ulcorner\phi\urcorner = \delta\in_A\ulcorner\phi\urcorner = \delta\phi = \phi,$$

so $\ulcorner\delta\in_A\urcorner\ulcorner\phi\urcorner = \ulcorner\phi\urcorner$ by 3.4.6(ii). Thus by 3.1.6 we have

$$\pi\ulcorner\phi\urcorner = (\ulcorner\delta\in_A\urcorner\cap 1_{PA})\ulcorner\phi\urcorner = \ulcorner\delta\in_A\urcorner\ulcorner\phi\urcorner\cap\ulcorner\phi\urcorner = \ulcorner\phi\urcorner.$$

Next, we have

$$\phi\gamma\ulcorner\phi\urcorner^\circ \leq \in_A\ulcorner\phi\urcorner\ulcorner\phi\urcorner^\circ = \in_A\pi\ulcorner\phi\urcorner\ulcorner\phi\urcorner^\circ \leq \in_A\pi,$$

so $\gamma\ulcorner\phi\urcorner^\circ \leq (\phi\backslash\in_A\pi)$. And $(\in_A\pi\backslash\phi) \geq (\in_A\backslash\phi) \geq \ulcorner\phi\urcorner$ since π is a core, so

$$\gamma(\phi\backslash\in_A\pi)\pi(\in_A\pi\backslash\phi)\gamma \geq \gamma\ulcorner\phi\urcorner^\circ\pi\ulcorner\phi\urcorner\gamma = \gamma\ulcorner\phi\urcorner^\circ\ulcorner\phi\urcorner\gamma \geq \gamma,$$

as required.

(ii) Let δ be an equivalence. Then we have

$$\ulcorner\delta\urcorner^\circ\ulcorner\delta\urcorner = (\delta|\in_A)(\in_A|\delta) \leq (\delta|\delta) = \delta,$$

but also

$$\delta = (\delta|\delta) \leq (\delta|\delta)\ulcorner\delta\urcorner^\circ\ulcorner\delta\urcorner = (\delta|\delta)(\delta|\in_A)\ulcorner\delta\urcorner \leq (\delta|\in_A)\ulcorner\delta\urcorner = \ulcorner\delta\urcorner^\circ\ulcorner\delta\urcorner,$$

so we may take f to be $\ulcorner\delta\urcorner$.

(iii) Again, let δ be an equivalence, and factor it as $f^\circ f$. Then ff° is a core, so we can split it (by 3.3.3(ii)) as gg° with $g^\circ g = 1$. Let $\phi = g^\circ f$; we have $\phi^\circ\phi = f^\circ gg^\circ f = f^\circ ff^\circ f = \delta\delta = \delta$ and $\phi\phi^\circ = g^\circ ff^\circ g = g^\circ gg^\circ g = 1$, so $\phi^\circ\phi$ is a splitting of δ (and ϕ is a map). The second assertion follows from (i) and 3.3.12.

(iv) Let A be an object of $\mathcal{A}$, and let 0_A denote the bottom element of $\mathcal{A}(A, A)$. Since composition in $\mathcal{A}$ preserves bottom elements, it is easy to see that $\tau_A = 0_A/0_A$ (or, if you prefer, $1_A/0_A$) is the top element of $\mathcal{A}(A, A)$, and hence that it is an equivalence. If we split it as $g^\circ g$ where $g\colon A \to U$ (using 3.3.3(iii)), then 1_U is the top element of $\mathcal{A}(U, U)$, since if $\phi\colon U \looparrowright U$ then

$$\phi = gg^\circ\phi gg^\circ \leq g\tau_A g^\circ = gg^\circ gg^\circ = 1_U\ .$$

So U will be a unit provided every object of $\mathcal{A}$ admits a map to it; but A admits a map to U, and if we take $A = PB$ for some B then every object admits a map to A, viz. the name of the zero morphism to B. □

Corollary 3.4.9 *If $\mathcal{A}$ is a nonempty pre-tabular power allegory whose hom-posets have bottom elements, then* $\mathbf{Map}(\mathcal{A}[\check{\mathcal{K}}])$ *is a topos.*

Proof By 3.4.8(i), $\mathcal{A}[\check{\mathcal{K}}]$ is a power allegory, and by 3.3.6(ii) it is tabular. By 3.4.8(iii) it is effective, and its hom-posets inherit bottom elements from $\mathcal{A}$; hence it has a unit by 3.4.8(iv). So the result follows from 3.4.7. □

We conclude this section with a non-elementary result which enables us to recognize certain particular allegories as being power allegories. First we need some new terminology. We call a morphism $\phi\colon A \looparrowright B$ in an allegory a *partial map* if it satisfies $\phi\phi^\circ \leq 1_B$, but not necessarily $\phi^\circ\phi \geq 1_A$. We shall use the notation $\phi\colon A \rightharpoondown B$ for partial maps; it is easy to see that in $\mathbf{Rel}(\mathcal{C})$ partial maps are exactly what we called by that name in Section A2.4. And we shall call an object G of an allegory a *strong separator* if, given any two morphisms $\phi, \psi\colon A \looparrowright B$ with $\phi \neq \psi$, there exists not merely a morphism but a partial map $\chi\colon G \rightharpoondown A$ with $\phi\chi \neq \psi\chi$.

Lemma 3.4.10 *If G is a strong separator in an allegory $\mathcal{A}$, then $I(G)$ is a strong separator in $\mathcal{A}[\check{\mathcal{E}}]$, for any class $\mathcal{E}$ of symmetric idempotents in $\mathcal{A}$.*

Proof Let $\phi, \psi : \epsilon \looparrowright \delta$ be two morphisms of $\mathcal{A}[\check{\mathcal{E}}]$ with $\phi \neq \psi$. Then there exists $\chi : G \rightharpoondown A$ in $\mathcal{A}$ (where $A = \mathrm{dom}\ \epsilon$) with $\phi\chi \neq \psi\chi$; now $\epsilon\chi$ is a morphism $I(G) \looparrowright \epsilon$ in $\mathcal{A}[\check{\mathcal{E}}]$ which has distinct composites with ϕ and ψ, and

$$(\epsilon\chi)(\epsilon\chi)^\circ = \epsilon\chi\chi^\circ\epsilon \leq \epsilon\epsilon = \epsilon,$$

so $\epsilon\chi$ is a partial map. □

Theorem 3.4.11 *Let $\mathcal{A}$ be an effective positive geometric allegory with a strong separator. Then $\mathcal{A}$ is a power allegory.*

Proof Let G be a strong separator, and let A be an arbitrary object of $\mathcal{A}$. Form the coproduct H of copies of G indexed by $\mathcal{A}(G, A)$, and let $\theta : H \looparrowright A$ be the unique morphism satisfying $\theta n_\phi = \phi$ for all $\phi : G \looparrowright A$ (here n_ϕ is, as usual, the ϕth coprojection). $(\theta|\theta)$ is an equivalence (recall that a geometric allegory is necessarily a division allegory), so we can split it as $g^\circ g$ for some $g : H \to K$; let $\epsilon = \theta g^\circ : K \looparrowright A$. We shall show that this latter morphism satisfies the conditions of 3.4.5.

First, we note that $\epsilon g = \theta g^\circ g \geq \theta$ but also

$$\epsilon g = \theta(\theta|\theta) \leq \theta(\theta\backslash\theta) \leq \theta,$$

so $\epsilon g = \theta$. Now

$$\theta g^\circ(\epsilon\backslash\epsilon)g = \epsilon(\epsilon\backslash\epsilon)g \leq \epsilon g = \theta,$$

so $g^\circ(\epsilon\backslash\epsilon)g \leq (\theta\backslash\theta)$, whence $g^\circ(\epsilon|\epsilon)g \leq (\theta|\theta)$ and so

$$(\epsilon|\epsilon) = gg^\circ(\epsilon|\epsilon)gg^\circ \leq g(\theta|\theta)g^\circ = gg^\circ gg^\circ = 1_K\,.$$

Next, suppose we have a factorization $\epsilon = \delta h$ where $h : K \to L$ is a map and $(\delta|\delta) = 1_L$. Since $h^\circ h$ is symmetric and $hh^\circ h = h$, we have $h^\circ h \leq (h|h) \leq (\epsilon|\epsilon) = 1_K$, so $h^\circ h = 1_K$. Suppose $hh^\circ \neq 1_L$; then there is a partial map $\chi : G \rightharpoondown L$ with $hh^\circ\chi \neq \chi$. But now $\delta\chi : G \looparrowright A$, so we can factor $\delta\chi$ as $\theta n_{\delta\chi} = \delta h g n_{\delta\chi} = \delta h f$, say, where f is a map. Thus $\delta \geq \delta\chi\chi^\circ = \delta h f \chi^\circ$, so $hf\chi^\circ \leq (\delta\backslash\delta)$; but similarly $\delta \geq \delta\chi f^\circ h^\circ$, so $\chi f^\circ h^\circ \leq (\delta\backslash\delta)$ and hence $hf\chi^\circ \leq (\delta|\delta) = 1_L$. Now we have

$$\chi \leq \chi f^\circ h^\circ h f \leq hf;$$

we claim that $\chi = hf(\chi^\circ\chi \cap 1_G)$. For we have

$$hf(\chi^\circ\chi \cap 1_G) \leq hf\chi^\circ\chi \leq \chi$$

and

$$hf(\chi^\circ\chi \cap 1_G) \geq \chi(\chi^\circ\chi \cap 1_G) \geq \chi \cap \chi = \chi\,.$$

So $hh^\circ\chi = hh^\circ hf(\chi^\circ\chi \cap 1_G) = hf(\chi^\circ\chi \cap 1_G) = \chi$, contradicting our assumption. Thus we have shown that, if there is a factorization $\epsilon = \delta h$ with $(\delta|\delta) = 1$, then necessarily h is an isomorphism.

Now let $\phi\colon B \looparrowright A$ be an arbitrary morphism of $\mathcal{A}$, and consider the morphism $\psi\colon B + K \looparrowright A$ defined by $\psi n_1 = \phi$, $\psi n_2 = \epsilon$. By splitting the equivalence $(\psi|\psi)$ as $k^\circ k$ and arguing as before, we can factor ψ as ωk, where $(\omega|\omega) = 1$. But then $\omega k n_1 = \epsilon$, so kn_1 is an isomorphism, and hence $\phi = \omega k n_2 = \epsilon(kn_1)^\circ kn_2$; that is, we can factor ϕ as ϵl where l is a map. Now $(\epsilon\backslash\phi) \geq l$, and also $\phi l^\circ = \epsilon l l^\circ \leq \epsilon$ whence $(\phi\backslash\epsilon) \geq l^\circ$, so

$$(\phi\backslash\epsilon)(\epsilon\backslash\phi) \geq l^\circ l \geq 1_B,$$

as required. □

Corollary 3.4.12 *If $\mathcal{A}$ is a pre-tabular positive geometric allegory with a strong separator, then* $\mathbf{Map}(\mathcal{A}[\check{S}])$ *is a topos.*

Proof Putting together all our results on splitting symmetric idempotents, we know that $\mathcal{A}[\check{S}]$ is tabular, effective and positive geometric, and that it has a strong separator. Hence by 3.4.11 it is a power allegory, and by 3.4.8(iv) it has a unit. So the result follows from 3.4.7. □

In particular, we may now prove a long-promised result:

Corollary 3.4.13 *For any frame L,* $\mathbf{Set}(L)$ *is a topos.*

Proof By definition, $\mathbf{Set}(L)$ is the category of maps of $\mathbf{Mat}(L)[\check{S}]$, and we know that $\mathbf{Mat}(L)$ is pre-tabular and positive geometric, so it is enough to show that $\mathbf{Mat}(L)$ has a strong separator. But any singleton set $\{*\}$ will do for this; for if ϕ and ψ are two distinct $B \times A$ matrices, they must differ at some particular entry, say $\phi(b_0, a_0) \neq \psi(b_0, a_0)$, and then the matrix

$$\begin{aligned} \chi(a,*) &= 1 && \text{if } a = a_0 \\ &= 0 && \text{otherwise} \end{aligned}$$

defines a map (not just a partial map) $\{*\} \to A$ satisfying $\phi\chi \neq \psi\chi$. □

We shall see in B3.1.9 that there is a very large class of toposes which may be constructed in this way (that is, by splitting all symmetric idempotents and then cutting down to the subcategory of maps) from relatively 'simple' positive geometric allegories having sets as objects, disjoint union as coproduct and a singleton set as strong separator.

Suggestions for further reading: Carboni et al. [225, 226], Freyd & Scedrov [381].

A4

GEOMETRIC MORPHISMS – BASIC THEORY

A4.1 Definition and examples

In Chapters A1 and A2 we took the view that, whenever we introduced a new class of categories, there should be a corresponding class of functors, namely those which preserve the structure involved in the definition of the given class of categories. Thus the notion of topos, when we introduced it, came with the associated notion of logical functor. However, an important aspect of topos theory, and one which particularly reflects the confluence of geometrical and logical influences in the origins of the subject, is the fact that it has another natural notion of 'morphism of toposes' associated with it: the notion of geometric morphism. We shall see in 4.1.18 below (and again in B2.2.7) that geometric morphisms are also 'structure-preserving functors' in a suitable sense, although the structure they preserve is different from that involved in our original definition of a topos.

In fact a geometric morphism is not a single functor but an adjoint pair of functors. One could, of course, define a geometric morphism to be a functor possessing an adjoint rather than an adjoint pair: since the adjoint of a functor, if it has one, is unique up to canonical natural isomorphism, this would not make any essential difference, but in practice it is generally convenient to have both the left and right adjoints explicitly specified as part of the definition of a geometric morphism, and we shall take this approach.

Definition 4.1.1 (*a*) Let $\mathcal{E}$ and $\mathcal{F}$ be toposes. A *geometric morphism* $f\colon \mathcal{F} \to \mathcal{E}$ consists of a pair of functors $f_*\colon \mathcal{F} \to \mathcal{E}$ (the *direct image* of f) and $f^*\colon \mathcal{E} \to \mathcal{F}$ (the *inverse image* of f) together with an adjunction $(f^* \dashv f_*)$, such that f^* is cartesian (i.e. preserves finite limits).

(*b*) Let f and $g\colon \mathcal{F} \to \mathcal{E}$ be geometric morphisms. A *geometric transformation* $\alpha\colon f \to g$ is defined to be a natural transformation $\alpha\colon f^* \to g^*$.

A few comments on this definition are in order. We do not have to demand, in part (*a*), that the direct image functor f_* should be cartesian, since this follows from its possession of a left adjoint – indeed, it preserves whatever limits exist in $\mathcal{F}$. Similarly, the inverse image f^* preserves whatever colimits exist in $\mathcal{E}$. The fact that a geometric morphism f is regarded as pointing in the same direction as f_* is a purely conventional choice; we shall see later that, in a sense, f_* 'embodies the geometric aspects' of the morphism f, and f^* embodies

its algebraic aspects. There is a similar arbitrary choice of direction involved in part (b) of the definition: in general, if we are given adjunctions $(f^* \dashv f_*)$ and $(g^* \dashv g_*)$ between the same two categories, then natural transformations $f^* \to g^*$ correspond bijectively to natural transformations $g_* \to f_*$. Both choices are justified on pragmatic grounds: they minimize the amount of dualization we shall have to do later on in the discussion of particular examples (see, for instance, 4.1.4 and 4.1.11 below).

It is also possible to consider geometric morphisms and transformations (defined in exactly the same way) between quasitoposes rather than toposes. At various points in this chapter, we shall prove results about geometric morphisms between quasitoposes rather than toposes (when we can do so without significant extra effort), but we shall not develop the full theory of such morphisms here.

We shall write $\mathfrak{Top}$ for the 2-category of toposes, geometric morphisms and geometric transformations. Although this is a 2-category, we shall generally treat it as if it were a bicategory: that is, we shall not expect diagrams of toposes and geometric morphisms to commute 'on the nose', but only up to (specified) invertible geometric transformations. For instance, when we refer (as we frequently shall) to the slice category $\mathfrak{Top}/\mathcal{E}$, we shall mean the 2-category whose objects are morphisms $f: \mathcal{F} \to \mathcal{E}$ in $\mathfrak{Top}$, whose 1-cells $(g: \mathcal{G} \to \mathcal{E}) \to (f: \mathcal{F} \to \mathcal{E})$ are pairs (h, α) where h is a geometric morphism $\mathcal{G} \to \mathcal{F}$ and α is a 2-isomorphism $g \to fh$, and whose 2-cells $(h, \alpha) \to (k, \beta)$ are geometric transformations $\gamma: h \to k$ such that $f\gamma \circ \alpha = \beta$. Normally, we shall take such details for granted, but we shall spell them out in full (as here) whenever there is a serious danger of confusion. (See also Section B1.1 for a more detailed discussion of our general attitude to diagrams in 2-categories.)

The primary justification for studying geometric morphisms is that examples of such morphisms occur 'all over the place' when we look at particular examples of toposes. We devote most of this section to introducing some of these examples.

Example 4.1.2 Let $f: A \to B$ be a morphism in a topos $\mathcal{E}$. By 2.3.2, the slice categories $\mathcal{E}/A$ and $\mathcal{E}/B$ are toposes, and by 2.3.3 the pullback functor $f^*: \mathcal{E}/B \to \mathcal{E}/A$ is the inverse image of a geometric morphism, whose direct image is the functor denoted Π_f in 1.5.3. In this way the assignment $A \mapsto \mathcal{E}/A$ becomes a functor $\mathcal{E} \to \mathfrak{Top}$ (but only in the 'pseudo' sense indicated above: if $g: B \to C$ is another morphism of $\mathcal{E}$, we do not expect to have an equality $(gf)^* = f^*g^*$, but only a canonical isomorphism between these functors). In fact, recalling the isomorphism $\mathcal{E} \cong \mathcal{E}/1$, we can regard $A \mapsto \mathcal{E}/A$ as a functor $\mathcal{E} \to \mathfrak{Top}/\mathcal{E}$.

The geometric morphisms which arise as in Example 4.1.2 are rather special, in that their inverse image functors are logical. In 2.3.8 above, we saw that this does not quite suffice to characterize them amongst geometric morphisms; in Section C3.5 we study what can be said in general about geometric morphisms

with logical inverse images. Incidentally, 2.3.9 shows that the only geometric morphisms with logical direct images are equivalences.

Example 4.1.3 Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and A any object of $\mathcal{E}$. Then we have a geometric morphism (which we shall denote by (f/A)) from the slice category $\mathcal{F}/f^*A$ to $\mathcal{E}/A$: explicitly, $(f/A)^*$ is simply f^* applied to morphisms of $\mathcal{E}$ with codomain A, and $(f/A)_*(g: B \to f^*A)$ is the left vertical morphism in the pullback

$$\begin{tikzcd} C \arrow[r] \arrow[d] & f_*B \arrow[d, "f_*(g)"] \\ A \arrow[r, "\eta_A"] & f_*f^*A \end{tikzcd} ,$$

where η is the unit of $(f^* \dashv f_*)$. It is straightforward to verify that $(f/A)^*$ is cartesian and left adjoint to $(f/A)_*$. Note also that, for any morphism $g: A \to B$ in $\mathcal{E}$, the square

$$\begin{tikzcd} \mathcal{F}/f^*A \arrow[r] \arrow[d, "(f/A)"] & \mathcal{F}/f^*B \arrow[d, "(f/B)"] \\ \mathcal{E}/A \arrow[r] & \mathcal{E}/B \end{tikzcd}$$

(where the horizontal morphisms are defined as in 4.1.2 from f^*g and g respectively) commutes up to natural isomorphism.

We shall say that a property of geometric morphisms is *stable under slicing* if it is inherited by morphisms of the form (f/A) from any morphism f which possesses it.

Example 4.1.4 Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories. Then composition with f defines a functor $f^*: [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$, which has adjoints on both sides, the left and right *Kan extensions* along f: for example, the right Kan extension $\varprojlim_f$ sends a functor $F: \mathcal{C} \to \mathbf{Set}$ to the functor whose value at an object B of $\mathcal{D}$ is the limit of the diagram

$$(B \downarrow f) \xrightarrow{U} \mathcal{C} \xrightarrow{F} \mathbf{Set}$$

(here $(B \downarrow f)$ is the comma category whose objects are pairs (A, ϕ) with $\phi: B \to fA$ in $\mathcal{D}$, and U is the forgetful functor from this category to $\mathcal{C}$). Thus f^* is the inverse image of a geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$, whose direct image is $\varprojlim_f$. Moreover, any natural transformation $\alpha: f \to g$ between functors $\mathcal{C} \to \mathcal{D}$ induces a natural transformation $f^* \to g^*$ (whose value at F is the natural transformation $F\alpha: Ff \to Fg$), i.e. a geometric transformation

$(\varprojlim_f, f^*) \to (\varprojlim_g, g^*)$. Thus the assignment $\mathcal{C} \mapsto [\mathcal{C}, \mathbf{Set}]$ can be made into a functor (that is, a 2-functor) from the 2-category $\mathfrak{Cat}$ of small categories, functors and natural transformations into $\mathfrak{Top}$ (in fact into $\mathfrak{Top}/\mathbf{Set}$).

The particular case of 4.1.4 in which the categories $\mathcal{C}$ and $\mathcal{D}$ are discrete corresponds, via the equivalence $\mathbf{Set}^B \simeq \mathbf{Set}/B$ of 1.1.6, to the particular case $\mathcal{E} = \mathbf{Set}$ of 4.1.2. In B2.3.22 we shall generalize Example 4.1.4, replacing $\mathbf{Set}$ by an arbitrary topos $\mathcal{E}$ (and the notion of small category by that of an internal category in $\mathcal{E}$); for the present, we may note that it admits an immediate generalization from $\mathbf{Set}$ to $\mathbf{Set}_f$ (taking $\mathcal{C}$ and $\mathcal{D}$ to be finite categories).

We note that the geometric morphisms which arise as in 4.1.4, though not as special as those of 4.1.2, still have the property that their inverse image functors have left adjoints as well as right adjoints. We call a geometric morphism f *essential* if it has this property; we normally write $f_!$ for the left adjoint of f^*. With the aid of this notion, we can prove a partial converse to 4.1.4:

Lemma 4.1.5 *Let $\mathcal{C}$ and $\mathcal{D}$ be small categories such that $\mathcal{D}$ is Cauchy-complete (cf.* 1.1.10*). Then every essential geometric morphism $f\colon [\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is induced as in* 4.1.4 *by a functor $\mathcal{C} \to \mathcal{D}$.*

Proof Since $f_!$ has a right adjoint f^* which preserves epimorphisms, it is straightforward to verify that $f_!$ preserves projectives; similarly, the fact that f^* preserves coproducts implies that $f_!$ preserves indecomposable objects. Hence by 1.1.10 $f_!$ maps representable functors in $[\mathcal{C}, \mathbf{Set}]$ to representables in $[\mathcal{D}, \mathbf{Set}]$, i.e. it restricts via the Yoneda embeddings to a functor $f_0\colon \mathcal{C} \to \mathcal{D}$. The adjunction $(f_! \dashv f^*)$, and the Yoneda lemma, then yield isomorphisms

$$\begin{aligned} f^*(F)(A) &\cong [\mathcal{C}, \mathbf{Set}]\,(\mathcal{C}\,(A, -), f^*(F)) \\ &\cong [\mathcal{D}, \mathbf{Set}]\,(\mathcal{D}\,(f_0(A), -), F) \\ &\cong F(f_0(A)) \end{aligned}$$

for every $F\colon \mathcal{D} \to \mathbf{Set}$ and every $A \in \mathrm{ob}\,\mathcal{C}$; so f^* is identified with the functor 'compose with f_0', and hence $f_!$ and f_* are respectively isomorphic to $\varinjlim_{f_0}$ and $\varprojlim_{f_0}$. □

We note that 4.1.5, like 4.1.4, can be made 2-functorial: if $\alpha\colon f \to g$ is a geometric transformation between essential geometric morphisms $[\mathcal{C}, \mathbf{Set}] \rightrightarrows [\mathcal{D}, \mathbf{Set}]$, then α induces a natural transformation $g_! \to f_!$, which restricts (since the Yoneda embeddings are contravariant) to a natural transformation $f_0 \to g_0$. In B3.2.8(c) we shall characterize arbitrary geometric morphisms between toposes of the form $[\mathcal{C}, \mathbf{Set}]$: they do not all arise as in 4.1.4, but they may all be constructed from certain 'generalized functors'. (Cf. also 4.1.10 below.)

We now return to our list of examples of geometric morphisms.

Example 4.1.6 Let G be a (nearly discrete) topological group. In Example 2.1.6 we noted that the inclusion $\mathbf{Cont}(G) \to [G, \mathbf{Set}]$ is cartesian

and has a right adjoint; so it is the inverse image of a geometric morphism $[G, \mathbf{Set}] \to \mathbf{Cont}(G)$.

Example 4.1.7 Let $f: G \to H$ be a continuous homomorphism between topological groups. If we regard f as a functor between small categories, then it induces a functor $f^*: [H, \mathbf{Set}] \to [G, \mathbf{Set}]$ as in 4.1.4; and it is easy to see that this functor maps continuous H-sets to continuous G-sets, since if B is an H-set then the stabilizers of points of f^*B are just the inverse images under f of their stabilizers as points of B. Thus we can regard f^* as a functor $\mathbf{Cont}(H) \to \mathbf{Cont}(G)$; as such, it is clearly still cartesian, and it has a right adjoint obtained by applying $\varprojlim_f$ and then coreflecting into $\mathbf{Cont}(H)$. In this way the assignment $G \mapsto \mathbf{Cont}(G)$ becomes a functor from the category $\mathbf{Gp}(\mathbf{Sp})$ of topological groups to $\mathfrak{Top}$.

Example 4.1.8 Let $(\mathcal{C}, T)$ be a small site, as defined in 2.1.9. The inclusion functor $\mathbf{Sh}(\mathcal{C}, T) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ has a cartesian left adjoint (the *associated sheaf functor* – this is a special case of a result which we shall prove in 4.4.8 below); so it is the direct image of a geometric morphism.

Example 4.1.9 By 1.4.7, there is (up to canonical isomorphism) exactly one cartesian and cocartesian functor $\Delta: \mathbf{Set}_f \to \mathcal{E}$ for any topos $\mathcal{E}$, namely the functor which sends an n-element set to the n-fold copower of 1. If it has a right adjoint Γ, then for any object A of $\mathcal{E}$ we have

$$\Gamma A \cong \mathbf{Set}_f(1, \Gamma A) \cong \mathcal{E}(\Delta 1, A) \cong \mathcal{E}(1, A);$$

taking the right-hand side as a definition, it is easy to verify that it does define a functor $\mathcal{E} \to \mathbf{Set}_f$ right adjoint to Δ, provided $\mathcal{E}(1, A)$ is finite for every object A of $\mathcal{E}$. But since $\mathcal{E}$ is cartesian closed, this is equivalent to requiring that every hom-set $\mathcal{E}(A, B) \cong \mathcal{E}(1, B^A)$ be finite. Thus we have shown that, *up to unique isomorphism, a topos admits at most one geometric morphism to* $\mathbf{Set}_f$, *and it admits such a morphism iff it is locally finite (i.e. has finite hom-sets).* A similar argument shows that *there is at most one geometric morphism* $\mathcal{E} \to \mathbf{Set}$ *for any* $\mathcal{E}$, *and such a morphism exists iff* $\mathcal{E}$ *is locally small and has arbitrary set-indexed copowers of* 1. (The latter condition does not imply that $\mathcal{E}$ has all small coproducts; compare Example 2.1.7.)

Example 4.1.10 Let $\mathcal{C}$ and $\mathcal{D}$ be small cartesian categories, and $f: \mathcal{C} \to \mathcal{D}$ a cartesian functor. We shall show that in this case the left Kan extension functor $\varinjlim_f: [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is also cartesian, so that it is the inverse image of a geometric morphism $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, whose *direct* image is f^* (compare 4.1.4). To verify this, note that for any $B \in \mathrm{ob}\, \mathcal{D}$, the functor $\varinjlim_f(-)(B): [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}$ may be described as the composite

$$[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{U^*} [(B \downarrow f)^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{\varinjlim} \mathbf{Set}$$

where $U: (B \downarrow f) \to \mathcal{C}$ is the forgetful functor, as before. Now U^* is cartesian, since it has a left adjoint, so the composite will be cartesian (by a well-known property of filtered colimits in **Set**) provided we can show that the category $(B \downarrow f)^{\mathrm{op}}$ is filtered. But this follows easily from the cartesianness of f: it is nonempty because there is a morphism $B \to 1 \cong f1$, and if $(\phi: B \to fA)$ and $(\phi': B \to fA')$ are any two objects of $(B \downarrow f)$, then

$$B \xrightarrow{(\phi,\phi')} fA \times fA' \cong f(A \times A')$$

is a third object which maps to both of them. Similarly, if ψ and $\psi': A \to A'$ both define morphisms in $(B \downarrow f)$ from $(\phi: B \to fA)$ to $(\phi': B \to fA')$, then ϕ factors through $f\psi''$, where $\psi'': A'' \to A$ is the equalizer of ψ and ψ'. Thus $(B \downarrow f)^{\mathrm{op}}$ is filtered, and so $\varinjlim_f(-)(B)$ is cartesian; but from the definition of limits in $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ this implies that $\varinjlim_f$ is cartesian. Note that, since we are here dealing with contravariant functors, a natural transformation $\alpha: f \to g$ between cartesian functors $\mathcal{C} \to \mathcal{D}$ induces a natural transformation $g^* \to f^*$ and hence a natural transformation $\varinjlim_f \to \varinjlim_g$; so, if we write $\mathfrak{Cart}$ for the 2-category of small cartesian categories, cartesian functors and natural transformations, the assignments $\mathcal{C} \mapsto [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ and $f \mapsto (f^*, \varinjlim_f)$ define a functor $\mathfrak{Cart}^{\mathrm{op}} \to \mathfrak{Top}$ (where the superscript 'op' indicates that 1-cells but not 2-cells are reversed).

Example 4.1.11 Let $f: X \to Y$ be a continuous map of topological spaces. Then we may regard f^{-1} as a functor $\mathcal{O}(Y) \to \mathcal{O}(X)$. From the fact that f^{-1} preserves arbitrary unions and finite intersections, we see that if $F: \mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Set}$ is a sheaf on X then the composite $F \circ f^{-1}: \mathcal{O}(Y)^{\mathrm{op}} \to \mathbf{Set}$ is a sheaf on Y, i.e. the functor $(f^{-1})^*: [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{O}(Y)^{\mathrm{op}}, \mathbf{Set}]$ maps the subcategory $\mathbf{Sh}(X)$ into $\mathbf{Sh}(Y)$. We write $f_*: \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ for the restriction of $(f^{-1})^*$. As a functor between presheaf categories, $(f^{-1})^*$ has a cartesian left adjoint $\varinjlim_{f^{-1}}$, by 4.1.10; this adjoint does not in general map sheaves to sheaves, but by composing it with the associated sheaf functor $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Sh}(X)$ (cf. 4.1.8) we obtain a cartesian left adjoint $f^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ for f_*, and hence a geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$. Once again, this construction is functorial – that is, $X \mapsto \mathbf{Sh}(X)$ defines a functor $\mathbf{Sp} \to \mathfrak{Top}$.

It was, of course, Example 4.1.11 which motivated both the definition and the name of geometric morphism. We shall return to it in C1.4.3, where we shall see that the inverse image f^* has an alternative (and simpler) description in terms of local homeomorphisms, and also that there is an appropriate notion of 2-cell in **Sp** to which the above functor can be extended; and in Section C2.3 we shall see how the example may be extended to a suitable notion of 'continuous functor' between more general sites.

Example 4.1.12 Let $F: \mathcal{E} \to \mathcal{F}$ be a cartesian functor between toposes, and let $\mathbf{Gl}(F)$ be the topos obtained by glueing along F, as in 2.1.12. The two projection functors $P_1: \mathbf{Gl}(F) \to \mathcal{E}$ and $P_2: \mathbf{Gl}(F) \to \mathcal{F}$ are both cartesian, and

both have right adjoints: the right adjoint R_1 of P_1 sends an object A of $\mathcal{E}$ to $(A, FA, 1_{FA})$, and the right adjoint R_2 of P_2 sends B to $(1, B, B \to 1 \cong F1)$. Thus we have geometric morphisms from both $\mathcal{E}$ and $\mathcal{F}$ to $\mathbf{Gl}(F)$. Note that the composite

$$\mathcal{E} \xrightarrow{R_1} \mathbf{Gl}(F) \xrightarrow{P_2} \mathcal{F}$$

is simply F; thus we have shown that *any cartesian functor between toposes can be factored as a direct image functor followed by an inverse image functor.* Note also that the functor P_1 is logical, by the constructions of 2.1.12 (although P_2 is not in general); in fact, as we shall see in 4.5.6 below, $\mathcal{E}$ is equivalent in $\mathfrak{Top}/\mathbf{Gl}(F)$ to the slice category $\mathbf{Gl}(F)/(1, 0, 0 \to 1)$.

Remark 4.1.13 Occasionally in this book (particularly in Part B), we shall wish to consider adjunctions between toposes of which the left adjoint preserves pullbacks (and hence all finite connected limits, by 1.2.9), but not the terminal object. We shall call these *pre-geometric morphisms*, and use the same notation for them as for geometric morphisms. (In [1041], they are called 'partial geometric morphisms'.) We note that if $f\colon \mathcal{F} \to \mathcal{E}$ is a pre-geometric morphism, then we may factor it as

$$\mathcal{F} \xrightarrow{g_I} \mathcal{F}/I \xrightarrow{\hat{f}} \mathcal{E}$$

where $I = f^*(1)$, g_I is the pre-geometric morphism defined by $g_I^* = \Sigma_I$ and $g_{I*} = I^*$, and $\hat{f}$ is a geometric morphism: specifically, $\hat{f}^*(A)$ is f^* applied to the unique morphism $A \to 1$, as in 1.2.9, and $\hat{f}_*(B \to I)$ is the pullback of

$$\begin{tikzcd} & f_*(B) \arrow[d] \\ 1 \arrow[r, "\eta_1"] & f_*(I) \end{tikzcd}$$

where η is the unit of $(f^* \dashv f_*)$. Thus results about pre-geometric morphisms can usually be deduced from the corresponding results about geometric morphisms and the particular case of morphisms of the form g_I.

We conclude this section with a number of unrelated results about geometric morphisms, which do not fit into the general development of the theory in the rest of this chapter, but which deserve to be mentioned at this stage. First we prove a simple but important result about geometric morphisms and natural number objects.

Lemma 4.1.14 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be a geometric morphism. If $\mathcal{E}$ has a natural number object, then so has $\mathcal{F}$.*

Proof If N is a natural number object in $\mathcal{E}$, then it follows from Lemma 2.5.6(i) that f^*N is a natural number object in $\mathcal{F}$. $\square$

The converse of 4.1.14 is not true in general: if $f: \mathcal{F} \to \mathcal{E}$ is a geometric morphism and $\mathcal{F}$ has a natural number object, there is no reason why $\mathcal{E}$ should have one. For a simple counterexample, take $\mathcal{F}$ to be **Set** and $\mathcal{E}$ to be the topos obtained by glueing along the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$ (which does not have a natural number object, since $\mathbf{Set}_f$ admits a geometric morphism to it). However, in the particular case when f is a surjection (which we shall study in the next section), then it will follow from 4.2.1(v) below that we can construct a natural number object for $\mathcal{E}$ from one for $\mathcal{F}$.

We have not said very much in this section (and will not say much in the rest of this chapter) about the 2-categorical structure of $\mathfrak{Top}$; detailed discussion of this aspect of topos theory is postponed until Part B. However, there is one simple result which might as well be proved now. In general, the hom-categories $\mathfrak{Top}(\mathcal{F}, \mathcal{E})$ are not well-endowed with limits or colimits, but they do have the very minimum of such structure:

Lemma 4.1.15 *$\mathfrak{Top}$ is locally Cauchy complete; that is, its idempotent 2-cells split.*

Proof Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism and $\epsilon: f \to f$ an idempotent geometric transformation. Then for each object A of $\mathcal{E}$ the morphism $\epsilon_A: f^*A \to f^*A$ is idempotent; since $\mathcal{F}$ has equalizers, we can split it as $f^*A \twoheadrightarrow g^*A \rightarrowtail f^*A$. From the naturality of ϵ, it is easy to see that g^* becomes a functor $\mathcal{E} \to \mathcal{F}$, and that it preserves finite limits. Now the mate $\bar{\epsilon}: f_* \to f_*$ of ϵ is also idempotent, and so we can similarly split $\bar{\epsilon}_B: f_*B \to f_*B$ as $f_*B \twoheadrightarrow g_*B \rightarrowtail f_*B$ for each object B of $\mathcal{F}$. Finally, morphisms $A \to g_*B$ correspond to morphisms $h: A \to f_*B$ satisfying $\bar{\epsilon}_B h = h$, and hence to morphisms $\bar{h}: f^*A \to B$ satisfying $\bar{h}\epsilon_A = \bar{h}$, which in turn correspond to morphisms $g^*A \to B$. So the pair (g_*, g^*) is a geometric morphism $\mathcal{F} \to \mathcal{E}$, which is clearly a splitting for ϵ in $\mathfrak{Top}(\mathcal{F}, \mathcal{E})$. □

It is also appropriate at this point to say something about Beck–Chevalley conditions in $\mathfrak{Top}$. Given a commutative square (up to isomorphism)

$$\begin{tikzcd}
\mathcal{H} \arrow[r, "k"] \arrow[d, "h"'] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"'] & \mathcal{E}
\end{tikzcd}$$

of toposes and geometric morphisms, the natural isomorphism $k^*g^* \cong h^*f^*$ transposes to yield a canonical natural transformation $\theta: g^*f_* \to k_*h^*$ between functors $\mathcal{F} \rightrightarrows \mathcal{G}$. We say that the square *satisfies the Beck–Chevalley condition* if θ is an isomorphism; we shall also be interested in the condition that it is pointwise monic (i.e. θ_B is monic for each object B of $\mathcal{F}$), which we shall

call the *weak Beck–Chevalley condition.* (In practice, it is usual to consider the Beck–Chevalley condition only for pullback squares in $\mathfrak{Top}$; but the definition, and the two preliminary results to be proved here, do not require the assumption that the square is a pullback.) We shall have a good deal to say about Beck–Chevalley conditions in Chapter C3, but there are a couple of results about them which do not involve any of the machinery of Part C, and so might as well be given now. The first is a simple stability property:

Lemma 4.1.16 *Suppose*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

satisfies the (weak) Beck–Chevalley condition. Then so does the square

$$\begin{tikzcd} \mathcal{H}/h^*f^*A \arrow[r, "k/g^*A"] \arrow[d, "h/f^*A"] & \mathcal{G}/g^*A \arrow[d, "g/A"] \\ \mathcal{F}/f^*A \arrow[r, "f/A"] & \mathcal{E}/A \end{tikzcd}$$

for any object A of $\mathcal{E}$.

Proof Let $p\colon B \to f^*A$ be an object of $\mathcal{F}/f^*A$. We have pullback squares

$$\begin{tikzcd} g^*C \arrow[r] \arrow[d] & g^*f_*B \arrow[d] \\ g^*A \arrow[r] & g^*f_*f^*A \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} D \arrow[r] \arrow[d] & k_*h^*B \arrow[d] \\ g^*A \arrow[r] & k_*k^*g^*A \cong k_*h^*f^*A \end{tikzcd}$$

(the former since g^* preserves pullbacks) whose left vertical morphisms are respectively the images of p under $(g/A)^*(f/A)_*$ and under $(k/g^*A)_*(h/f^*A)^*$. A straightforward diagram-chase shows that if the comparison maps between the right-hand sides of the squares are isomorphisms (respectively, monic), then the same is true of the induced map between their top left corners. $\square$

The second result is more substantial – and more surprising: it says (approximately) that the weak Beck–Chevalley condition is equivalent to the restriction of the strong one to subterminal objects.

Proposition 4.1.17 *For a commutative square*

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{k} & \mathcal{G} \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle g} \\ \mathcal{F} & \xrightarrow{f} & \mathcal{E} \end{array}$$

in $\mathfrak{Top}$, the following are equivalent:

(i) *The square satisfies the weak Beck–Chevalley condition.*

(ii) *For any monomorphism $B' \rightarrowtail B$ in $\mathcal{F}$, the naturality square*

$$\begin{array}{ccc} g^* f_* B' & \xrightarrow{\theta_{B'}} & k_* h^* B' \\ \downarrow & & \downarrow \\ g^* f_* B & \xrightarrow{\theta_B} & k_* h^* B \end{array}$$

is a pullback.

(iii) *For each object A of $\mathcal{E}$, the square*

$$\begin{array}{ccc} \mathcal{H}/h^* f^* A & \xrightarrow{k/g^* A} & \mathcal{G}/g^* A \\ {\scriptstyle h/f^* A}\downarrow & & \downarrow{\scriptstyle g/A} \\ \mathcal{F}/f^* A & \xrightarrow{f/A} & \mathcal{E}/A \end{array}$$

satisfies the Beck–Chevalley condition for subterminal objects, i.e. the natural transformation (θ/A) associated with this square has the property that $(\theta/A)_m$ is an isomorphism for each subterminal object $m\colon B \rightarrowtail f^* A$ of $\mathcal{F}/f^* A$.

Proof (i)$\Leftrightarrow$(ii) follows from 1.6.9, since the functors $g^* f_*$ and $k_* h^*$ are cartesian. 1.6.9 also shows that these conditions imply the particular case $A = 1$ of (iii); but we have seen in 4.1.16 that condition (i) is stable under slicing, and hence it implies the general case of (iii). So it remains to prove that (iii) implies (ii).

Given a monomorphism $m\colon B' \rightarrowtail B$ in $\mathcal{F}$, set $A = f_*B$ and form the pullback

$$\begin{tikzcd}
C \arrow[r] \arrow[d, tail, "n"] & B' \arrow[d, tail, "m"] \\
f^*A \arrow[r, "\epsilon_B"] & B
\end{tikzcd}$$

where ϵ is the counit of $(f^* \dashv f_*)$. The triangular identities for the adjunction, and the fact that f_* preserves pullbacks, tell us that $(f/A)_*(n)$ is simply $f_*(m)\colon f_*B' \rightarrowtail f_*B$; so $(g/A)^*(f/A)_*(n)$ is just $g^*f_*(m)$. But $(k/g^*A)_*(h/f^*A)^*(n)$ is by definition the left vertical morphism in the diagram

$$\begin{tikzcd}
D \arrow[r] \arrow[d, tail] & k_*h^*C \arrow[r] \arrow[d, tail, "k_*h^*(n)"] & k_*h^*B' \arrow[d, tail, "k_*h^*(m)"] \\
g^*A \arrow[r] & k_*k^*g^*A \cong k_*h^*f^*A \arrow[r] & k_*h^*B
\end{tikzcd}$$

where both squares are pullbacks, and the first factor of the bottom edge is the unit of $(k^* \dashv k_*)$. Now a straightforward diagram-chase shows that the bottom composite is simply θ_B, and that $(\theta/A)_n$ is the factorization through the pullback of $\theta_{B'}\colon g^*f_*B' \to k_*h^*B'$. So the assertion that $(\theta/A)_n$ is an isomorphism implies that the naturality square for θ at m is a pullback. □

Finally in this section, we outline one sense (not the only one) in which geometric morphisms – or at least inverse image functors – can be viewed as the structure-preserving maps between toposes: namely the allegorical sense. If $f\colon \mathcal{F} \to \mathcal{E}$ is a geometric morphism, then $f^*\colon \mathcal{E} \to \mathcal{F}$ is a regular functor (though f_* is not, in general), and so extends to a unital morphism of allegories which we shall denote $F\colon \mathbf{Rel}(\mathcal{E}) \to \mathbf{Rel}(\mathcal{F})$, by the remarks after 3.2.10. Further, since f^* preserves coproducts and images, F preserves unions of morphisms to the extent that they exist in $\mathbf{Rel}(\mathcal{E})$; that is, it is a morphism of union allegories (and, if the toposes concerned have arbitrary set-indexed unions of subobjects, then it is a morphism of geometric allegories). We cannot hope to obtain an elementary converse result; but if we restrict our attention to the class of allegories studied in 3.4.11, then we have

Proposition 4.1.18 *Let $\mathcal{A}$ and $\mathcal{B}$ be allegories satisfying the hypotheses of 3.4.11. Then unital morphisms of geometric allegories $\mathcal{A} \to \mathcal{B}$ correspond bijectively to inverse image functors* $\mathbf{Map}(\mathcal{A}) \to \mathbf{Map}(\mathcal{B})$.

Proof We have seen that an inverse image functor induces a unital morphism of geometric allegories. Conversely, suppose given such a morphism $F\colon \mathcal{A} \to \mathcal{B}$,

and let f^* denote its restriction to $\mathbf{Map}(\mathcal{A})$ (which of course takes values in $\mathbf{Map}(\mathcal{B})$, since F preserves adjunctions). Since F preserves arbitrary (set-indexed) unions, f^* preserves arbitrary unions of subobjects; hence, by an infinitary extension of the argument after 1.4.4, it preserves arbitrary coproducts. As a regular functor between effective regular categories, it also preserves coequalizers of equivalence relations; and it preserves the construction of the equivalence relation generated by a parallel pair $u, v \colon A \rightrightarrows B$ (that is, the smallest equivalence relation on B through which (u, v) factors – or equivalently the kernel-pair of the coequalizer of u and v), since the latter may be constructed as the image of a map to $B \times B$ from a countable coproduct of objects constructed from u and v by finite limits. So f^* preserves all coequalizers, and hence all set-indexed colimits.

Now if G is a strong separator for $\mathcal{A}$, as defined before 3.4.10, then it is readily seen that the subobjects of G in $\mathbf{Map}(\mathcal{A})$ form a separating set for $\mathbf{Map}(\mathcal{A})$ (they form a set because $\mathbf{Map}(\mathcal{A})$, as a subcategory of $\mathcal{A}$, is locally small, and it has a subobject classifier). So we are in a position to apply the Special Adjoint Functor Theorem (cf. C2.2.10) to obtain a right adjoint f_* for f^*; thus f^* is the inverse image of a geometric morphism $\mathbf{Map}(\mathcal{B}) \to \mathbf{Map}(\mathcal{A})$. □

Two cautionary remarks about the foregoing result should be made. First, the reader should not be deceived into thinking that 4.1.18 represents the origin of the term 'geometric morphism'; rather, the term 'geometric allegory' was chosen in order to fit with it. Secondly, while 4.1.18 may be viewed as 'explaining' the notion of geometric morphism, it does not provide a similarly satisfactory explanation of the notion of geometric transformation; as we remarked after 3.2.10, natural transformations between regular functors do not correspond to any very obvious notion of '2-cell' between the corresponding morphisms of allegories.

Suggestions for further reading: Carboni et al. [225, 226], Rosebrugh & Wood [1041], Wraith [1236].

A4.2 Surjections and inclusions

A geometric morphism $f \colon \mathcal{F} \to \mathcal{E}$, like any other adjunction, induces a monad on one of the categories involved (in this case $\mathcal{E}$) and a comonad on the other. We shall defer discussion of the monads induced by geometric morphisms until later; in this section our concern is with the comonads.

We note first that the functor part of such a comonad must be cartesian, since it is the composite of the direct and inverse images of f. Of course, we shall use the term *cartesian comonad* for a comonad with this property. We shall see shortly that every cartesian comonad on a topos arises from a geometric morphism. First, we prove an omnibus theorem which indicates how much categorical structure is inherited by the category of coalgebras for such a comonad.

Theorem 4.2.1 *Let $\mathcal{E}$ be a cartesian category, and $\mathbb{G} = (G, \epsilon, \delta)$ a cartesian comonad on $\mathcal{E}$. Let $\mathcal{E}_{\mathbb{G}}$ denote the category of $\mathbb{G}$-coalgebras in $\mathcal{E}$.*

(i) *If $\mathcal{E}$ is properly cartesian closed, so is $\mathcal{E}_{\mathbb{G}}$.*
(ii) *If $\mathcal{E}$ is locally cartesian closed, so is $\mathcal{E}_{\mathbb{G}}$.*
(iii) *If $\mathcal{E}$ has a subobject classifier, so has $\mathcal{E}_{\mathbb{G}}$.*
(iv) *If $\mathcal{E}$ is coregular and has a weak subobject classifier, then $\mathcal{E}_{\mathbb{G}}$ has a weak subobject classifier.*
(v) *If $\mathcal{E}$ has a natural number object, so has $\mathcal{E}_{\mathbb{G}}$.*

In particular,

(vi) *If $\mathcal{E}$ is a topos (resp. a quasitopos), so is $\mathcal{E}_{\mathbb{G}}$.*

Proof (i) $\mathcal{E}_{\mathbb{G}}$ is cartesian, since the forgetful functor $\mathcal{E}_{\mathbb{G}} \to \mathcal{E}$ creates finite limits. Now let (A, α) and (B, β) be $\mathbb{G}$-coalgebras, and let $E \rightarrowtail G(B^A)$ be the equalizer of

$$\begin{tikzcd} G(B^A) \arrow[rr, "G(\beta^A)"] \arrow[dr, "\delta_{B^A}"'] & & G(GB^A) \\ & GG(B^A) \arrow[r, "G\theta"] & G(GB^{GA}) \arrow[u, "G(GB^\alpha)"'] \end{tikzcd}$$

where $\theta\colon G(B^A) \to GB^{GA}$ is the exponential comparison map for G, i.e. the transpose of

$$G(B^A) \times GA \cong G(B^A \times A) \xrightarrow{G(\mathrm{ev})} GB\,.$$

If we equip the vertices of the above diagram with cofree coalgebra structures, then the edges are all coalgebra homomorphisms; so E acquires a coalgebra structure which we shall denote $\eta\colon E \to GE$. Now let (C, γ) be another coalgebra; then coalgebra homomorphisms $(C, \gamma) \to (G(B^A), \delta_{B^A})$ correspond bijectively to morphisms $C \to B^A$ in $\mathcal{E}$, and hence to morphisms $C \times A \to B$. But the statement that $f\colon C \times A \to B$ is a coalgebra homomorphism $(C, \gamma) \times (A, \alpha) \to (B, \beta)$ is equivalent to saying that the diagram

$$\begin{tikzcd} C \arrow[r, "\overline{f}"] \arrow[d, "\gamma"] & B^A \arrow[r, "\beta^A"] & GB^A \\ GC \arrow[r, "G\overline{f}"] & G(B^A) \arrow[r, "\theta"] & GB^{GA} \arrow[u, "GB^\alpha"'] \end{tikzcd}$$

commutes, which in turn is equivalent to saying that the corresponding coalgebra homomorphism $(C,\gamma) \to (G(B^A),\delta)$ factors (uniquely) through the subcoalgebra (E,η). So (E,η) has the universal property of an exponential $(B,\beta)^{(A,\alpha)}$ in $\mathcal{E}_{\mathbb{G}}$.

(ii) Let (A,α) be a $\mathbb{G}$-coalgebra. To show that $\mathcal{E}_{\mathbb{G}}/(A,\alpha)$ is cartesian closed, we shall show that it is isomorphic to $(\mathcal{E}/A)_{\mathbb{G}'}$ for a suitable cartesian comonad $\mathbb{G}'$ on $\mathcal{E}/A$, and then appeal to (i). To define G' at an object $f\colon B \to A$ of $\mathcal{E}/A$, we form the pullback

$$\begin{tikzcd} G'B \arrow[r, "h"] \arrow[d, "G'f"] & GB \arrow[d, "Gf"] \\ A \arrow[r, "\alpha"] & GA \end{tikzcd} \quad ;$$

this is clearly functorial. Moreover, since $\epsilon_A\alpha = 1_A$, the composite $\epsilon_B h\colon G'B \to GB \to B$ is a morphism over A, and thus defines a natural transformation ϵ' from G' to the identity. And since G preserves pullbacks, $G'G'B$ is the pullback of GGB along the composite $G\alpha \circ \alpha\colon A \to GGA$; but this composite equals $\delta_A \circ \alpha$, and so the morphism δ_B induces a morphism $\delta'_f\colon G'B \to G'G'B$ over A, which is also natural in f. It is straightforward to verify that ϵ' and δ' form a comonad structure on G'. Moreover, this comonad is cartesian: G' preserves pullbacks since G and the pullback functor $\alpha^*\colon \mathcal{E}/GA \to \mathcal{E}/A$ both do so, and it preserves the terminal object by construction. Finally, for any $f\colon B \to A$, there is a bijective correspondence between morphisms $f \to G'f$ in $\mathcal{E}/A$ and morphisms $\beta\colon B \to GB$ satisfying $Gf \circ \beta = \alpha \circ f$; and it is again easy to see that this restricts to a correspondence between $\mathbb{G}'$-coalgebra structures on f and $\mathbb{G}$-coalgebra structures on B which make f a coalgebra homomorphism.

(iii) Suppose $\mathcal{E}$ has a subobject classifier Ω. Form the equalizer $\Omega_{\mathbb{G}} \rightarrowtail G\Omega$ of

$$\begin{tikzcd} G\Omega \arrow[rr, "1"] \arrow[dr, "\delta_\Omega"'] & & G\Omega \\ & GG\Omega \arrow[ur, "G\tau"'] & \end{tikzcd}$$

where τ is the classifying map of $G(\top)\colon 1 \cong G1 \to G\Omega$; as in (i), we may equip $\Omega_{\mathbb{G}}$ with a coalgebra structure by regarding the vertices of this diagram as cofree coalgebras. Now, given a $\mathbb{G}$-coalgebra (A,α), the subobjects of (A,α) in $\mathcal{E}_{\mathbb{G}}$ are just those subobjects $A' \rightarrowtail A$ of A in $\mathcal{E}$ for which the composite of $A' \rightarrowtail A$ with α admits a (necessarily unique) factorization α' through $GA' \rightarrowtail GA$; but if such

a factorization exists then the square

$$\begin{tikzcd} A' \arrow[r, "\alpha'"] \arrow[d, tail] & GA' \arrow[d, tail] \\ A \arrow[r, "\alpha"] & GA \end{tikzcd}$$

is necessarily a pullback, since α and α' are split by ϵ_A and $\epsilon_{A'}$ respectively. If $\phi\colon A \to \Omega$ is the classifying map of $A' \rightarrowtail A$, then the composite $\tau \circ G\phi$ classifies $GA' \rightarrowtail GA$; so, by the above remarks, $A' \rightarrowtail A$ is a subcoalgebra of A iff the diagram

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[d, "\alpha"] & \Omega \\ GA \arrow[r, "G\phi"] & G\Omega \arrow[u, "\tau"'] \end{tikzcd}$$

commutes. But this in turn is equivalent to saying that the coalgebra homomorphism $(A, \alpha) \to (G\Omega, \delta_\Omega)$ corresponding to ϕ factors through the equalizer $\Omega_{\mathbb{G}}$ defined above; so the latter has the universal property of a subobject classifier for $\mathcal{E}_{\mathbb{G}}$.

(iv) The argument for weak subobject classifiers is identical, except that we need to show that cocovers in $\mathcal{E}_{\mathbb{G}}$ are precisely those coalgebra homomorphisms $(A, \alpha) \rightarrowtail (B, \beta)$ which are cocovers in $\mathcal{E}$. One direction is easy, since the forgetful functor $\mathcal{E}_{\mathbb{G}} \to \mathcal{E}$ preserves epimorphisms and reflects isomorphisms. Conversely, suppose $f\colon (A, \alpha) \to (B, \beta)$ is a cocover in $\mathcal{E}_{\mathbb{G}}$. Form its epic-cocover factorization

$$A \xrightarrow{e} \!\!\!\!\rightarrow A' \xrightarrow{f'} B$$

in $\mathcal{E}$; then since cocovers in $\mathcal{E}$ coincide with regular monomorphisms, Gf' is a cocover, and hence we obtain a unique $\alpha'\colon A' \to GA'$ making

$$\begin{tikzcd} A \arrow[r, two heads, "e"] \arrow[d, "\alpha"] & A' \arrow[r, tail, "f'"] \arrow[d, "\alpha'"] & B \arrow[d, "\beta"] \\ GA \arrow[r, two heads, "Ge"] & GA' \arrow[r, tail, "Gf'"] & GB \end{tikzcd}$$

commute. It is now clear that (A', α') is a $\mathbb{G}$-coalgebra, and that $e\colon (A, \alpha) \to (A', \alpha')$ is an epimorphism in $\mathcal{E}_{\mathbb{G}}$ through which f factors; so e is an isomorphism, and f is a cocover in $\mathcal{E}$, as required.

(v) Suppose (N, o, s) is a natural number object in $\mathcal{E}$, and let $\nu\colon N \to GN$ be the unique morphism making

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[d, equal] & N \arrow[r, "s"] \arrow[d, "\nu"] & N \arrow[d, "\nu"] \\ G1 \arrow[r, "Go"] & GN \arrow[r, "Gs"] & GN \end{tikzcd}$$

commute. It follows easily from the uniqueness clause of Definition 2.5.1 that ν satisfies the commutative diagrams required for a $\mathbb{G}$-coalgebra structure on N. Moreover, if (A, α) is a $\mathbb{G}$-coalgebra, and $a\colon 1 \to (A, \alpha)$ and $t\colon (A, \alpha) \to (A, \alpha)$ are coalgebra homomorphisms, then a similar application of the uniqueness clause shows that the unique $f\colon N \to A$ making

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "a"'] & N \arrow[r, "s"] \arrow[d, "f"] & N \arrow[d, "f"] \\ & A \arrow[r, "t"] & A \end{tikzcd}$$

commute is a coalgebra homomorphism $(N, \nu) \to (A, \alpha)$. So $((N, \nu), o, s)$ is a natural number object in $\mathcal{E}_{\mathbb{G}}$.

(vi) For toposes, this is immediate from (i) and (iii); for quasitoposes, it follows from (ii), (iv) and the fact that $\mathcal{E}_{\mathbb{G}}$ inherits finite colimits from $\mathcal{E}$. □

Corollary 4.2.2 *For any cartesian comonad $\mathbb{G}$ on a topos $\mathcal{E}$, there is a geometric morphism $f\colon \mathcal{E} \to \mathcal{F}$ such that $\mathbb{G}$ is (isomorphic to) the comonad induced by $(f^* \dashv f_*)$.*

Proof Take $\mathcal{F}$ to be $\mathcal{E}_{\mathbb{G}}$, and let f be the geometric morphism whose direct and inverse images are the cofree and forgetful functors. □

Remark 4.2.3 With the notion of pre-geometric morphism, as introduced in 4.1.13, in mind, we might ask what happens if we relax the cartesianness condition in 4.2.2 to the mere preservation of pullbacks (equivalently, by 1.2.9, of finite connected limits). The answer is that the result survives essentially unchanged. Let $\mathbb{G} = (G, \epsilon, \delta)$ be a comonad on a topos $\mathcal{E}$, such that G preserves pullbacks. Then, applying the construction of 4.2.1(ii) to the terminal $\mathbb{G}$-coalgebra $(G1, \delta_1)$, and noting that this construction only requires G to preserve pullbacks, we may construct a cartesian comonad $\mathbb{G}'$ on $\mathcal{E}/G1$, whose category of coalgebras is isomorphic to that of $\mathbb{G}$; so this category is a topos, and the cofree and forgetful functors form a pre-geometric morphism $\mathcal{E} \to \mathcal{E}_{\mathbb{G}}$.

We have already seen a number of examples of toposes which are comonadic over other toposes. We list a few of them here, together with a couple of new examples (of which the second is of considerable importance).

Examples 4.2.4 (*a*) Let G be a topological group. Then the category $\mathbf{Cont}(G)$ of Example 2.1.6 is coreflective in $[G, \mathbf{Set}]$ and closed under finite limits; so it is comonadic over $[G, \mathbf{Set}]$ by an idempotent cartesian comonad. Thus 4.2.1 provides a new proof that $\mathbf{Cont}(G)$ is a topos. Note that in the case when the comonad $\mathbb{G}$ is idempotent (i.e. δ is an isomorphism), the construction in the proof of 4.2.1(i) may be appreciably simplified: the exponential $(B, \beta)^{(A,\alpha)}$ is just $G(B^A)$, since all the morphisms in the diagram defining E are isomorphisms. However, $\Omega_{\mathbb{G}}$ is not simply $G(\Omega)$, unless we make a further assumption about G – see 4.6.6 below.

(*b*) Let M be a monoid. Then the functor $(-)^M : \mathbf{Set} \to \mathbf{Set}$ has a comonad structure $\mathbb{M}$: the counit $\epsilon_A : A^M \to A$ sends a function $f : M \to A$ to $f(e)$, where e is the identity element of M, and the comultiplication δ_A similarly sends f to the function $x \mapsto (y \mapsto f(yx))$. Moreover, $(-)^M$ is cartesian, because it is right adjoint to $(-) \times M$. But it is easy to see that, for a set A, a function $A \to A^M$ is an $\mathbb{M}$-coalgebra structure on A iff its transpose $A \times M \to A$ is a right action of M on A in the usual sense; thus $\mathbf{Set}_{\mathbb{M}}$ is isomorphic to the category $[M^{\mathrm{op}}, \mathbf{Set}]$, and 4.2.1 provides a new proof that this functor category is a topos. More generally, we shall see in B2.3.17 that for any small category $\mathcal{C}$ the topos $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ may be identified with a category comonadic over the slice category $\mathbf{Set}/\mathrm{ob}\,\mathcal{C}$.

(*c*) Let $F : \mathcal{E} \to \mathcal{F}$ be a cartesian functor between toposes. We may define a cartesian comonad (G, ϵ, δ) on the cartesian product $\mathcal{E} \times \mathcal{F}$ as follows: the functor G sends (A, B) to $(A, B \times FA)$, the counit $\epsilon_{(A,B)}$ is $(1_A, \pi_1) : (A, B \times FA) \to (A, B)$, and the comultiplication δ is similarly induced by the diagonal map $FA \to FA \times FA$. It is easy to verify that a coalgebra structure for this comonad, on an object (A, B) of $\mathcal{E} \times \mathcal{F}$, is essentially the same thing as a morphism $B \to FA$, and hence that the category $(\mathcal{E} \times \mathcal{F})_{\mathbb{G}}$ is isomorphic to the category $\mathbf{Gl}(F)$ of 2.1.12. Since a cartesian product of toposes is trivially a topos, 4.2.1 thus provides a new proof that $\mathbf{Gl}(F)$ is a topos; and if we take $\mathcal{E}$ and $\mathcal{F}$ to be quasitoposes rather than toposes, then it establishes the result, promised in 2.6.7, that the glueing construction works for quasitoposes too. By 4.2.3, it also shows that $\mathbf{Gl}(F)$ is still a topos if ($\mathcal{E}$ and $\mathcal{F}$ are toposes and) we merely assume that F preserves pullbacks; but this could easily have been verified directly, since in this case $\mathbf{Gl}(F)$ is isomorphic to $\mathbf{Gl}(\hat{F})$, where $\hat{F} : \mathcal{E} \to \mathcal{F}/F1$ is the cartesian functor obtained by factoring F as in 1.2.9. (Rather remarkably, this result is best possible: if $F : \mathcal{E} \to \mathcal{F}$ is a functor between toposes such that $\mathbf{Gl}(F)$ is a topos, then F preserves pullbacks. See [224] for a proof.)

(*d*) Let $\mathcal{E}$ be a cocomplete, locally small topos with a separating set of objects (that is, a Grothendieck topos as defined in Section C2.2). The full subcategory $\mathcal{E}_d$ of decidable objects of $\mathcal{E}$ is closed under finite limits and arbitrary coproducts (this was proved in 1.4.15 except for the infinite coproducts, for which the proof

is just like that for finite ones); from this it follows that the full subcategory $\mathcal{E}_{qd}$ of quotients of decidable objects (i.e. objects A for which there exists an epimorphism $B \twoheadrightarrow A$ with B decidable) is closed under finite limits and arbitrary colimits. Thus we may apply the Special Adjoint Functor Theorem to obtain a right adjoint for the inclusion functor $\mathcal{E}_{qd} \to \mathcal{E}$: explicitly, the adjoint sends an object A to the union of all those subobjects of A which are quotients of decidable objects. Thus we have an idempotent comonad on $\mathcal{E}$ whose category of coalgebras is isomorphic to $\mathcal{E}_{qd}$, and it is cartesian because $\mathcal{E}_{qd}$ is closed under finite limits. So we may deduce that $\mathcal{E}_{qd}$ is a topos, for any such $\mathcal{E}$.

We remark that the use of non-elementary means (i.e. the Adjoint Functor Theorem) in this example cannot be avoided; for an arbitrary topos $\mathcal{E}$, $\mathcal{E}_{qd}$ may not be coreflective in $\mathcal{E}$. An example is provided by the topos $\mathcal{E}$ obtained by glueing along the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$ (cf. 2.1.12). It is easy to see that an object $(A, B, f\colon B \to A)$ of this topos is decidable iff f is injective (cf. 1.4.16), which forces B as well as A to be finite, and hence that $\mathcal{E}_{qd}$ is simply the category $[\mathbf{2}, \mathbf{Set}_f]$ whose objects are maps between finite sets. This is clearly not coreflective in $\mathcal{E}$ (although, by 2.1.3, it is a topos). We shall see in Section F2.2 that if $\mathcal{E}$ is the effective topos **Eff**, then $\mathcal{E}_{qd}$ is not even a topos. (On the other hand, if $\mathcal{E}$ admits a geometric morphism to a Boolean topos $\mathcal{S}$ which is bounded in the sense defined in B3.1.7, then we may use the $\mathcal{S}$-indexed version of the Adjoint Functor Theorem, proved in B2.4.6, to deduce that $\mathcal{E}_{qd}$ is coreflective in $\mathcal{E}$, and hence a topos; cf. [521]. In fact, neither of the counterexamples just mentioned admits *any* geometric morphism to a Boolean topos; see 4.5.24 below.)

(*e*) Let X be a topological space. We shall describe a cartesian comonad $\mathbb{G}$ on the topos $\mathbf{Set}/X$, whose category of coalgebras is isomorphic to $\mathbf{LH}/X$; this will prove that the latter is a topos, justifying claims made in Sections A1.5 and A1.6. Given an object $p\colon E \to X$ of $\mathbf{Set}/X$ and a point $x \in X$, we define a *local section* of p at x to be a function $s\colon U \to E$, where U is an open neighbourhood of x in X, such that ps is the inclusion $U \to X$. By a *germ of sections* of p at x we mean an equivalence class of local sections, where two sections $s\colon U \to E$ and $s'\colon U' \to E$ are considered equivalent if they agree on some open set V with $x \in V \subseteq U \cap U'$; and we write GE_x for the set of all germs of sections of p at x. Let GE denote the disjoint union of the GE_x, $x \in X$, and $Gp\colon GE \to X$ the obvious projection; then it is straightforward to verify that the assignment $p \mapsto Gp$ defines a cartesian functor $\mathbf{Set}/X \to \mathbf{Set}/X$. Moreover, we have a natural transformation ϵ from G to the identity functor, such that ϵ_p sends a germ of sections of p at x to its value at x (defined as the value at x of any local section representing the given germ; this is clearly independent of the choice of representative). And we have a natural transformation $\delta\colon G \to GG$: if $[s]_x$ is a germ of sections of p at x (represented by a local section $s\colon U \to E$, say), then $\delta_p([s]_x)$ is the germ at x of the local section $U \to GE$ whose value at $y \in U$ is $[s]_y$ (again, it is easy to verify that this germ is independent of the choice of s). It is also straightforward to check that (G, ϵ, δ) is a comonad on $\mathbf{Set}/X$. Now suppose that $((p\colon E \to X), (\alpha\colon E \to GE))$ is a $\mathbb{G}$-coalgebra. Then we may define

a topology on E, by specifying that $V \subseteq E$ is open iff, whenever $e \in V$, there is a representative $s: U \to E$ of the germ $\alpha(e)$ whose image is contained in V. This is clearly a topology, and it makes p a continuous map $E \to X$; but in fact it makes it a local homeomorphism since, for any $e \in E$, the restriction of p to the image of a representative of $\alpha(e)$ is a homeomorphism from an open set in E to an open set in X. Conversely, if we are given a topology on E such that p is a local homeomorphism, then we may define a $\mathbb{G}$-coalgebra structure α on p by setting $\alpha(e)$ to be the germ at $p(e)$ of any *continuous* local section of p passing through e. (The fact that such sections exist, and that they all have the same germ at $p(e)$, follows from the assumption that p is a local homeomorphism.) It is now not hard to check that these two constructions are inverse to each other, and that a map $f: (p: E \to X) \to (q: F \to X)$ between $\mathbb{G}$-coalgebras in $\mathbf{Set}/X$ is a coalgebra homomorphism iff it is continuous for the corresponding topologies on E and F. Thus we have established an isomorphism $(\mathbf{Set}/X)_{\mathbb{G}} \cong \mathbf{LH}/X$, as claimed.

Further important applications of Theorem 4.2.1 will be found in 4.6.3 and B2.3.17.

The following characterization of geometric morphisms between toposes of coalgebras is sometimes useful.

Proposition 4.2.5 *Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and let $\mathbb{G} = (G, \epsilon, \delta)$ and $\mathbb{H} = (H, \eta, \zeta)$ be cartesian comonads on $\mathcal{E}$ and $\mathcal{F}$ respectively. Then there is a bijection between (isomorphism classes of) geometric morphisms $g: \mathcal{F}_{\mathbb{H}} \to \mathcal{E}_{\mathbb{G}}$ making*

$$\begin{tikzcd}
\mathcal{F} \arrow[r, "f"] \arrow[d] & \mathcal{E} \arrow[d] \\
\mathcal{F}_{\mathbb{H}} \arrow[r, "g"] & \mathcal{E}_{\mathbb{G}}
\end{tikzcd}$$

commute (where the vertical arrows are the geometric morphisms defined as in 4.2.2), and natural transformations $\theta: f^*G \to Hf^*$ which are compatible with the comonad structures, in the sense that the diagrams

$$\begin{tikzcd}
f^*G \arrow[r, "\theta"] \arrow[dr, "f^*\epsilon"'] & Hf^* \arrow[d, "\eta_{f^*}"] \\
& f^*
\end{tikzcd}$$

and

$$\begin{tikzcd} f^*G \arrow[rr, "\theta"] \arrow[d, "f^*\delta"] & & Hf^* \arrow[d, "\zeta_{f^*}"] \\ f^*GG \arrow[r, "\theta_G"] & Hf^*G \arrow[r, "H\theta"] & HHf^* \end{tikzcd}$$

commute.

Proof It is well known that natural transformations θ satisfying the above commutative diagrams correspond to 'liftings' of f^* to functors $g^*: \mathcal{E}_{\mathbb{G}} \to \mathcal{F}_{\mathbb{H}}$ such that

$$\begin{tikzcd} \mathcal{E}_{\mathbb{G}} \arrow[r, "g^*"] \arrow[d] & \mathcal{F}_{\mathbb{H}} \arrow[d] \\ \mathcal{E} \arrow[r, "f^*"] & \mathcal{F} \end{tikzcd}$$

commutes (where the vertical arrows are the forgetful functors): the correspondence sends θ to the functor defined by

$$g^*(A, \alpha) = (f^*A, \theta_A \circ f^*\alpha),$$

and g^* to the transformation whose value at A is

$$f^*GA \xrightarrow{\beta} Hf^*GA \xrightarrow{Hf^*\epsilon_A} Hf^*A$$

where β is the structure map of the $\mathbb{H}$-coalgebra $g^*(GA, \delta_A)$. However, any such lifting g^* will be cartesian, since the forgetful functor $\mathcal{F}_{\mathbb{H}} \to \mathcal{F}$ creates finite limits; and since $\mathcal{E}_{\mathbb{G}}$ has equalizers, it will also inherit a right adjoint g_* from f_* by (the dual of) 1.1.3(i). Thus we have the desired correspondence between geometric morphisms and natural transformations θ as above. □

In particular, taking f to be the identity morphism on $\mathcal{E}$, we see that if $\mathbb{G}$ and $\mathbb{H}$ are two cartesian comonads on $\mathcal{E}$, then geometric morphisms $\mathcal{E}_{\mathbb{H}} \to \mathcal{E}_{\mathbb{G}}$ under $\mathcal{E}$ correspond to morphisms of comonads $\mathbb{G} \to \mathbb{H}$ in $\mathcal{E}$.

Next, we consider the class of geometric morphisms which arise (up to equivalence) as in 4.2.2.

Lemma 4.2.6 *Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism. The following conditions are equivalent:*

(i) *$\mathcal{E}$ is comonadic over $\mathcal{F}$; i.e. the comparison functor from $\mathcal{E}$ to the category of coalgebras for the comonad on $\mathcal{F}$ induced by $(f^* \dashv f_*)$ is (part of) an equivalence.*

(ii) *f^* is conservative.*
(iii) *f^* is faithful.*
(iv) *The unit η of the adjunction $(f^* \dashv f_*)$ is monic.*
(v) *For every monomorphism $m\colon A' \rightarrowtail A$ in $\mathcal{E}$, the naturality square*

$$\begin{tikzcd} A' \arrow[r, "\eta_{A'}"] \arrow[d, tail, "m"'] & f_*f^*A' \arrow[d, tail, "f_*f^*m"] \\ A \arrow[r, "\eta_A"] & f_*f^*A \end{tikzcd}$$

is a pullback.

Proof (i) $\Leftrightarrow$ (ii) follows from 1.1.2 (bearing in mind that f^* always preserves equalizers); (ii) $\Leftrightarrow$ (iii) holds since toposes are balanced categories; (iii) $\Leftrightarrow$ (iv) is a well-known property of adjunctions; and (iv) $\Leftrightarrow$ (v) is a particular case of 1.6.9. □

A geometric morphism satisfying the equivalent conditions of Lemma 4.2.6 is called a *surjection*. We next list some typical examples.

Examples 4.2.7 (*a*) Let $f\colon A \to B$ be a morphism in a topos $\mathcal{E}$. The geometric morphism $\mathcal{E}/A \to \mathcal{E}/B$ induced by f (as in 4.1.2) is a surjection iff f is an epimorphism; for we saw in 1.3.2(iii) that, in any regular category, the pullback functor f^* is conservative iff f is a cover.

(*b*) Let $f\colon \mathcal{C} \to \mathcal{D}$ be a functor between small categories. If f is surjective on objects, then it is easily verified that the functor $f^*\colon [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ is conservative; for a natural transformation α between functors $\mathcal{D} \to \mathbf{Set}$ is an isomorphism iff α_B is bijective for every object B of $\mathcal{D}$. So the geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ induced by f as in 4.1.4 is surjective. The converse is false; in fact, for the surjectivity of this morphism, it is necessary and sufficient that every object of $\mathcal{D}$ should be a retract of one in the image of f. The sufficiency is clear, since if B is a retract of fA then the component at B of any natural transformation between functors $\mathcal{D} \to \mathbf{Set}$ is uniquely determined by its component at fA. For the converse implication, suppose B is not a retract of any fA; let $F\colon \mathcal{D} \to \mathbf{Set}$ be the functor $\mathcal{D}(B, -)$ and $F' \rightarrowtail F$ the subfunctor defined by

$$F'(C) = \{u\colon B \to C \mid u \text{ is not split monic}\}\ .$$

Then the inclusion $F' \rightarrowtail F$ is not an isomorphism, but is mapped by f^* to an isomorphism.

(*c*) Let $f\colon X \to Y$ be a continuous map of topological spaces. If f is surjective, then the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by f as in 4.1.11 is a surjection. The simplest proof of this involves the identification of $\mathbf{Sh}(X)$ with

$\mathbf{LH}/X$, so we shall defer giving it until C1.5.1(ii). Note, however, that once we have this identification the surjection of 4.2.4(e) may be identified with that induced by the identity map $X_d \to X$ (where X_d denotes the set X re-topologized with the discrete topology), since $\mathbf{Set}/X$ is isomorphic to $\mathbf{LH}/X_d$.

Now let $f : \mathcal{F} \to \mathcal{E}$ be an arbitrary geometric morphism. Write $\mathbb{G}$ for the comonad on $\mathcal{F}$ induced by $(f^* \dashv f_*)$, and $g : \mathcal{F} \to \mathcal{F}_{\mathbb{G}}$ for the surjection of 4.2.2. The comparison functor $h^* : \mathcal{E} \to \mathcal{F}_{\mathbb{G}}$ satisfies $g^* h^* = f^*$, and is therefore cartesian since g^* creates finite limits. Moreover, since $\mathcal{E}$ has equalizers, h^* has a right adjoint h_*, which sends a $\mathbb{G}$-coalgebra (B, β) to the equalizer of

$$f_*B \underset{\eta_{f_*B}}{\overset{f_*\beta}{\rightrightarrows}} f_*f^*f_*B$$

(where η is the unit of $(f^* \dashv f_*)$), and $h_*g_* \cong f_*$ by the uniqueness of adjoints (or by direct calculation, if you prefer). Thus we have a factorization $f \cong hg$ of our original geometric morphism.

Proposition 4.2.8 *With the notation established above, the counit $h^*h_* \to 1$ is an isomorphism.*

Proof Let (B, β) be a $\mathbb{G}$-coalgebra. Since f^* preserves equalizers, we have an equalizer diagram

$$f^*h_*(B,\beta) \rightarrowtail f^*f_*B \underset{\delta_B}{\overset{f^*f_*\beta}{\rightrightarrows}} f^*f_*f^*f_*B$$

where $\delta = f^*\eta_{f_*}$ is the comultiplication of $\mathbb{G}$. But, by the definition of a $\mathbb{G}$-coalgebra, we also have a (split) equalizer diagram

$$B \overset{\beta}{\rightarrowtail} f^*f_*B \underset{\delta_B}{\overset{f^*f_*\beta}{\rightrightarrows}} f^*f_*f^*f_*B\,.$$

So there is a canonical isomorphism $f^*h_*(B,\beta) \to B$; but this isomorphism is easily seen to be the result of applying the forgetful functor g^* to the counit $h^*h_*(B,\beta) \to (B,\beta)$, and g^* reflects isomorphisms. □

A geometric morphism h satisfying the condition that the counit $h^*h_* \to 1$ is an isomorphism, or the equivalent condition that h_* is full and faithful, is called an *inclusion* (though some authors prefer the term *embedding*). We shall study inclusions in greater detail in the next three sections; for the present, we digress briefly to note an alternative characterization of them:

Lemma 4.2.9 *A geometric morphism is an inclusion iff its direct image is a cartesian closed functor (i.e. preserves exponentials).*

Proof Since inverse image functors preserve finite products, this is immediate from 1.5.9. □

We are now ready for the second main theorem of this section (the first being 4.2.1).

Theorem 4.2.10 *Every geometric morphism can be factored, uniquely up to canonical equivalence, as a surjection followed by an inclusion.*

Proof The existence of such a factorization has been established by Proposition 4.2.8 and the argument preceding it; so we have only to prove the uniqueness. Let

$$\begin{tikzcd} \mathcal{F} \arrow[rr, "f"] \arrow[dr, "\overline{g}"'] & & \mathcal{E} \\ & \mathcal{G} \arrow[ur, "\overline{h}"'] & \end{tikzcd}$$

be any factorization of a given morphism f into a surjection followed by an inclusion. Since the comonad on $\mathcal{G}$ induced by $(\overline{h}^* \dashv \overline{h}_*)$ is trivial, the comonad on $\mathcal{F}$ induced by $(\overline{g}^* \dashv \overline{g}_*)$ is isomorphic to that ($\mathbb{G}$, say) induced by $(f^* \dashv f_*)$. But $\overline{g}^*$ is comonadic by 4.2.6(i); i.e. there is an equivalence $\mathcal{G} \simeq \mathcal{F}_{\mathbb{G}}$ making

$$\begin{tikzcd} \mathcal{F} \arrow[r, "\overline{g}"] \arrow[dr, "g"'] & \mathcal{G} \arrow[d] \\ & \mathcal{F}_{\mathbb{G}} \end{tikzcd}$$

commute up to isomorphism. The fact that $\overline{g}^*$ reflects isomorphisms then shows that

$$\begin{tikzcd} \mathcal{G} \arrow[r, "\overline{h}"] \arrow[d] & \mathcal{E} \\ \mathcal{F}_{\mathbb{G}} \arrow[ur, "h"'] & \end{tikzcd}$$

commutes up to isomorphism. □

Corollary 4.2.11 *A geometric morphism which is both a surjection and an inclusion is an equivalence.* □

The corollary can easily be proved directly: if f is both a surjection and an inclusion, and η and ϵ denote the unit and counit of $(f^* \dashv f_*)$, then $f^*\eta$ is an isomorphism since it is a (one-sided) inverse for ϵ_{f^*}, but f^* reflects isomorphisms, so η is an isomorphism.

To conclude the present section, we revisit the three examples discussed in 4.2.7, and consider them in the light of 4.2.10.

Examples 4.2.12 (*a*) Let $f: A \to B$ be a morphism in a topos $\mathcal{E}$. As we saw in the proof of 2.4.7, the counit of the adjunction $(f^* \dashv \Pi_f)$ is an isomorphism iff the unit of $(\Sigma_f \dashv f^*)$ is an isomorphism; but the latter happens iff $f^*(f)$ is an isomorphism, i.e. iff f is monic. Combining this with 4.2.7(*a*), we see that the factorization of 4.2.10, applied to the morphism $\mathcal{E}/A \to \mathcal{E}/B$ induced by an arbitrary morphism $f: A \to B$ of $\mathcal{E}$, corresponds to the usual image factorization of f arising from the fact that $\mathcal{E}$ is a regular category.

(*b*) Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories. If f is full and faithful, then the induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is an *inclusion*; for the comma category $(fA \downarrow f)$ has an initial object 1_{fA} for any object A of $\mathcal{C}$, whence we deduce that $(\varprojlim_f(F))(fA) \cong F(A)$ for any $F: \mathcal{C} \to \mathbf{Set}$, i.e. $f^* \varprojlim_f(F) \cong F$. Thus, for an arbitrary f, the factorization of 4.2.10 applied to the morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ induced by f yields $[\mathcal{D}', \mathbf{Set}]$ as the image, where $\mathcal{D}'$ is the full image of f, i.e. the full subcategory of $\mathcal{D}$ on all objects of the form fA. Alternatively, by the last part of 4.2.7(*b*), we could 'complete' this subcategory by closing it off under images of idempotents in $\mathcal{D}$; of course, by 1.1.9, this makes no difference to the functor category $[\mathcal{D}', \mathbf{Set}]$.

(*c*) Let $f: X \to Y$ be a continuous map of topological spaces. Then it is straightforward to verify that $f_*: \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ is faithful iff it is full and faithful, iff $f^{-1}: \mathcal{O}(Y) \to \mathcal{O}(X)$ is surjective. If X is a subspace of Y and f is the inclusion, then the latter condition is satisfied; the converse holds (up to homeomorphism) provided Y satisfies the T_0 separation axiom, in which case the surjectivity of f^{-1} forces f to be injective. Combining this with 4.2.7(c), we see that if we apply the factorization of 4.2.10 to the morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by an arbitrary continuous $f: X \to Y$, we obtain $\mathbf{Sh}(I)$, where I is the image of f topologized as a subspace of Y (that is, we obtain the coimage factorization in $\mathbf{Sp}$, rather than the image factorization).

Suggestions for further reading: Carboni & Johnstone [224], Johnstone et al. [546].

A4.3 Cartesian reflectors and sheaves

In contrast to Theorem 4.2.1, the category of algebras for a cartesian monad $\mathbb{T} = (T, \eta, \mu)$ on a topos $\mathcal{E}$ is not in general a topos. (It can be shown to be cartesian closed, by an argument similar to that in the proof of 4.2.1(i): if (A, α) and (B, β) are $\mathbb{T}$-algebras, we define the exponential $(B, \beta)^{(A,\alpha)}$ to be the equalizer

of

$$\begin{tikzcd} B^A \arrow[rr, "B^\alpha"] \arrow[dr, "\eta_{B^A}"'] & & B^{TA} \\ & T(B^A) \xrightarrow{\theta} TB^{TA} \arrow[ur, "\beta^{TA}"'] & \end{tikzcd}$$

equipped with a suitable algebra structure. However, it lacks a subobject classifier in general.) But in the particular case when the monad is idempotent (i.e. μ is an isomorphism), so that the category of algebras may be identified with a reflective subcategory of $\mathcal{E}$, we do get a topos, as we shall see in this section. (We recall from Section A1.1 our conventions regarding the terms 'reflective subcategory', 'reflection' and 'reflector'. Normally we shall not distinguish notationally between a reflector (that is, an idempotent monad) and its functor part. We should also mention the tradition, deriving from module theory, whereby reflective subcategories of a cartesian category $\mathcal{C}$ for which the reflector is cartesian are called *localizations* of $\mathcal{C}$; but we shall not use this term.)

We begin with a simple but important result, which is a close relative of Lemma 1.5.8.

Proposition 4.3.1 *Let $\mathcal{E}$ be a cartesian closed category, and $\mathcal{L}$ a reflective subcategory of $\mathcal{E}$, corresponding to a reflector L on $\mathcal{E}$. Then L preserves finite products iff (the class of objects of) $\mathcal{L}$ is an exponential ideal in $\mathcal{E}$. Moreover, if these conditions hold then $B^\eta\colon B^{LA} \to B^A$ is an isomorphism for every object B of $\mathcal{L}$ and every object A of $\mathcal{E}$, where $\eta\colon 1_{\mathcal{E}} \to L$ is the unit of the reflection.*

Proof Suppose L preserves finite products; let B be an object of $\mathcal{L}$ and A an object of $\mathcal{E}$. For any object C of $\mathcal{E}$, we have the following string of bijective correspondences, which are natural in C:

$$\begin{array}{rcl} C & \longrightarrow & B^A \\ \hline C \times A & \longrightarrow & B \\ \hline LC \times LA & \longrightarrow & B \\ \hline LC \times A & \longrightarrow & B \\ \hline LC & \longrightarrow & B^A \end{array}$$

where we have used the fact that $LC \times LA$ is the reflection in $\mathcal{L}$ of both $C \times A$ and $LC \times A$. So every morphism $C \to B^A$ factors uniquely through $\eta_C\colon C \to LC$; applying this to the identity morphism $B^A \to B^A$, we deduce that η_{B^A} is an isomorphism, i.e. that B^A is in $\mathcal{L}$. Moreover, examination of the above argument shows that we have a natural bijection between morphisms $C \to B^A$ and $C \to B^{LA}$, so B^A and B^{LA} must be isomorphic (the isomorphism being B^η).

Conversely, suppose $\mathcal{L}$ is an exponential ideal. It is trivial that any reflector preserves 1, so we need only consider binary products. Let A and B be two

objects of $\mathcal{E}$; then for any object C of $\mathcal{L}$ we have natural bijections

$$\begin{array}{rcl} A \times B & \longrightarrow & C \\ \hline A & \longrightarrow & C^B \\ \hline LA & \longrightarrow & C^B \\ \hline LA \times B & \longrightarrow & C \\ \hline B & \longrightarrow & C^{LA} \\ \hline LB & \longrightarrow & C^{LA} \\ \hline LA \times LB & \longrightarrow & C, \end{array}$$

where we have used the fact that C^B and C^{LA} belong to $\mathcal{L}$ if C does. So every morphism from $A \times B$ to an object of $\mathcal{L}$ factors uniquely through $\eta_A \times \eta_B \colon A \times B \to LA \times LB$; but $LA \times LB$ belongs to $\mathcal{L}$, since the latter is closed under any limits which exist in $\mathcal{E}$, and so it must be (isomorphic to) the reflection $L(A \times B)$ of $A \times B$. □

Now suppose that $\mathcal{E}$ has pullbacks, and let L be a reflector on $\mathcal{E}$ which preserves pullbacks. Then, for any object A of $\mathcal{E}$, we may define a unary operation $c_{L,A}$ (or simply c_L) on subobjects of A, as follows: if $A' \rightarrowtail A$ is monic, then so is $LA' \rightarrowtail LA$, and we define $c_L(A')$ by the pullback diagram

$$\begin{tikzcd} c_L(A') \arrow[r] \arrow[d, tail] & LA' \arrow[d, tail] \\ A \arrow[r, "\eta_A"] & LA \end{tikzcd}$$

Lemma 4.3.2 *The operation c_L just defined is a closure operation on* Sub(A)*; that is, it is order-preserving and satisfies $A' \leq c_L(A') \cong c_L c_L(A')$ for any A'. Moreover, c_L commutes (up to isomorphism) with pullback along an arbitrary morphism of $\mathcal{E}$.*

Proof The fact that c_L is order-preserving is clear from the form of the definition, and the inequality $A' \leq c_L(A')$ follows immediately from the naturality of η. For the idempotency, we note that the pullback above is preserved by L; but $L\eta_A$ is an isomorphism, and so $Lc_L(A') \cong LLA' \cong LA'$ as subobjects of LA, whence $c_L c_L(A') \cong c_L(A')$. Finally, given a morphism $f \colon B \to A$, let $B' \rightarrowtail B$

be the pullback of $A' \rightarrowtail A$ along f; then in the diagram

$$\begin{tikzcd}
c_L(B') \arrow[rr] \arrow[dd, tail] \arrow[dr] & & LB' \arrow[dd] \arrow[dr] & \\
& c_L(A') \arrow[rr] \arrow[dd, tail] & & LA' \arrow[dd, tail] \\
B \arrow[rr] \arrow[dr] & & LB \arrow[dr] & \\
& A \arrow[rr] & & LA
\end{tikzcd}$$

the front, back and right vertical faces are pullbacks, whence it follows that the left one is too. $\square$

A closure operation on subobjects which commutes with pullback, as above, will be called a *universal closure operation* on $\mathcal{E}$. We shall use the terms *dense* and *closed* with their usual meanings relative to a universal closure operation. For future reference, we note some elementary properties of dense and closed subobjects:

Lemma 4.3.3 *Let* c *be a universal closure operation on a cartesian category* $\mathcal{E}$. *Then*

(i) *Given a commutative square*

$$\begin{tikzcd}
A' \arrow[r, "f'"] \arrow[d, tail, "m"] & B' \arrow[d, tail, "n"] \\
A \arrow[r, "f"] & B
\end{tikzcd}$$

where m *is a dense subobject and* n *is closed, there is a unique morphism* $g\colon A \to B'$ *satisfying* $ng = f$ *and* $gm = f'$.

(ii) *For any* $A' \rightarrowtail A$, $c(A')$ *may be characterized as the unique subobject* A'' *of* A *such that* $A' \rightarrowtail A''$ *is dense and* $A'' \rightarrowtail A$ *is closed.*

(iii) *For subobjects* $A' \rightarrowtail A$ *and* $A'' \rightarrowtail A$, *we have* $c(A' \cap A'') \cong c(A') \cap c(A'')$.

Proof (i) The commutativity of the square implies that $A' \leq f^*(B')$ in $\mathrm{Sub}(A)$. So we have

$$A \cong c(A') \leq c(f^*(B')) \cong f^*(c(B')) \cong f^*(B'),$$

which establishes the existence of g. Uniqueness follows from the fact that n is monic.

(ii) Clearly, $c(A') \rightarrowtail A$ is closed; and $A' \rightarrowtail c(A')$ is dense, since its closure is the pullback of $c(A') \rightarrowtail A$ along $c(A') \rightarrowtail A$. The fact that this characterizes $c(A')$ up to isomorphism in $\mathrm{Sub}(A)$ follows from two applications of (i) (in the particular case when f and f' are also monic).

(iii) By (ii), it suffices to show that $A' \cap A'' \rightarrowtail c(A') \cap c(A'')$ is dense and $c(A') \cap c(A'') \rightarrowtail A$ is closed. But this follows at once from the facts that the classes of closed and dense monomorphisms are both stable under pullback and composition, which in turn follow easily from the definition of a universal closure operation. □

Definition 4.3.4 Let c be a universal closure operation on a cartesian category $\mathcal{E}$.

(a) We say an object A of $\mathcal{E}$ is (c-)*separated* if, whenever we have a diagram

$$\begin{tikzcd} B' \arrow[r, "f'"] \arrow[d, tail, "m"] & A \\ B & \end{tikzcd}$$

where m is c-dense, there is at most one $f \colon B \to A$ with $fm = f'$.

(b) We say A is a (c-)*sheaf* if, whenever we have a diagram as above, there is exactly one f with $fm = f'$.

The following example explains the re-use of the terms which we met before in Chapter A2.

Example 4.3.5 Let $(\mathcal{C}, T)$ be a site, as defined in 2.1.9. Given a monomorphism $F' \rightarrowtail F$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ (where, as usual, we assume for convenience that each $F'(A)$ is an actual subset of $F(A)$), we define a new subfunctor $d(F') \rightarrowtail F$ by declaring that $d(F')(A)$ should contain, in addition to the elements of $F'(A)$, all those $x \in F(A)$ for which there exists a T-covering family $(f_i \colon A_i \to A \mid i \in I)$ such that $F(f_i)(x) \in F'(A_i)$ for each $i \in I$ (that is, those elements of $F(A)$ which are 'T-locally in F''). The stability condition (C) in the definition of a coverage ensures that $d(F')$ is indeed a functor; also, $d(F')$ contains F' by definition, and d is order-preserving as an operation on $\mathrm{Sub}(F)$. Unfortunately, d is not in general idempotent (though it is so in many particular cases, for example that

which arises from the usual notion of open covering on the open-set lattice of a topological space); but since $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is well-powered we may obtain a closure operation c by iterating d (transfinitely if necessary, taking unions at limit ordinals) until it converges. It is also easy to see that d commutes with pullback along a fixed morphism of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; hence so does its transfinite iterate, i.e. the closure operation c just described is universal.

We claim that the notions of T-sheaf and T-separated functor, which we defined in 2.1.9 and 2.6.4(d) respectively, coincide with those of c-sheaf and c-separated object defined in 4.3.4. In one direction, this is easy: given a T-covering family $(f_i \colon A_i \to A \mid i \in I)$, we form a subobject R of $\mathcal{C}(-, A)$ by setting

$$R(B) = \{g \colon B \to A \mid g \text{ factors through } f_i \text{ for some } i\} .$$

It is easy to verify that $R \rightarrowtail \mathcal{C}(-, A)$ is c-dense; in fact $d(R)$ is already the whole of $\mathcal{C}(-, A)$. Moreover, for any $F \colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, natural transformations $R \to F$ correspond bijectively to compatible families $s_i \in F(A_i)$ $(i \in I)$, as defined in 2.1.9; the correspondence sends $\theta \colon R \to F$ to the family $(\theta_{A_i}(f_i) \mid i \in I)$. So the assertion that F satisfies the sheaf axiom (resp. the separatedness axiom) for the family $(f_i \mid i \in I)$ says precisely that each morphism $R \to F$ has exactly one (resp. at most one) extension to a morphism $\mathcal{C}(-, A) \to F$. Thus c-sheaves (resp. c-separated objects) are T-sheaves (resp. T-separated functors).

For the converse, let F be a T-sheaf, $G' \rightarrowtail G$ a c-dense monomorphism, and $\theta \colon G' \to F$ a natural transformation. For any $x \in d(G')(A)$, either $x \in G'(A)$ or we have a covering $(f_i \colon A_i \to A \mid i \in I)$ such that each $G(f_i)(x)$ is in $G'(A_i)$; and then the family of elements

$$(\theta_{A_i}(G(f_i)(x)) \in F(A_i) \mid i \in I)$$

is clearly compatible and so defines a unique $y = d(\theta)_A(x) \in F(A)$. In this way we can extend θ uniquely to a natural transformation $d(\theta) \colon d(G') \to F$; repeating this process as often as necessary, we get a unique $c(\theta) \colon G \cong c(G') \to F$ extending θ. The argument for separated objects is similar.

The argument above does not establish a bijection between coverages on a small category $\mathcal{C}$ and universal closure operations on $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; several different coverages can give rise to the same closure operation (and consequently to the same category of sheaves). In Section C2.1 we shall distinguish a particular class of coverages (called *Grothendieck coverages*) which do correspond bijectively to universal closure operations on $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ (and to cartesian reflectors on $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$).

In the case where a universal closure operation is derived (as in 4.3.2) from a cartesian reflector, it is easy to recover the reflector (or at least the corresponding reflective subcategory) from it.

Lemma 4.3.6 *Let L be a cartesian reflector on a cartesian category $\mathcal{E}$, corresponding to a reflective subcategory $\mathcal{L}$, and let c_L denote the universal closure*

operation derived from L as in 4.3.2. Let A be an object of $\mathcal{E}$. Then

(a) The following are equivalent:

(i) *A is c_L-separated.*
(ii) *The unit map $\eta_A : A \to LA$ is monic.*
(iii) *A is a subobject of an object of $\mathcal{L}$.*
(iv) *The diagonal map $A \rightarrowtail A \times A$ is c_L-closed.*

(b) The following are equivalent:

(i) *A is a c_L-sheaf.*
(ii) *The unit $\eta_A : A \to LA$ is an isomorphism.*
(iii) *A is an object of $\mathcal{L}$.*

Proof (*a*) Assume (i); let $(a, b): R \rightrightarrows A$ be the kernel-pair of η_A, and $d: A \rightarrowtail R$ the diagonal map (witnessing the fact that R is reflexive), so that d is the equalizer of a and b. Since L is idempotent, $L\eta_A$ is an isomorphism; so since L preserves kernel-pairs and equalizers we have $La = Lb$ and Ld is an isomorphism. But this implies that d is c_L-dense; since A is separated we deduce that $a = b$, i.e. that η_A is monic. The implication (ii) $\Rightarrow$ (iii) is trivial. Condition (iv) holds for all objects of $\mathcal{L}$, since it is easy to see that all monomorphisms in $\mathcal{L}$ are c_L-closed; but it is inherited by subobjects, since the pullback of the diagonal along a monomorphism is the diagonal, and so (iii) implies (iv). Finally, (iv) implies (i) by an application of 4.3.3(i): if f_1 and f_2 are two morphisms completing the diagram of 4.3.4(*a*), then the diagram

$$\begin{tikzcd} B' \arrow[r, "f'"] \arrow[d, tail, "m"] & A \arrow[d, tail, "\Delta"] \\ B \arrow[r, "(f_1, f_2)"] & A \times A \end{tikzcd}$$

commutes, so (f_1, f_2) factors through the diagonal, i.e. $f_1 = f_2$.

(*b*) Again, let us begin by assuming (i). By part (*a*), η_A is monic, and since $L\eta_A$ is an isomorphism it is c_L-dense. So the identity morphism 1_A extends uniquely to a morphism $r: LA \to A$ with $r\eta_A = 1_A$. Thus A is a retract of an object of $\mathcal{L}$; but $\mathcal{L}$ is closed under retracts in $\mathcal{E}$, since it is closed under equalizers. So (iii) holds; the equivalence of (ii) and (iii) is again trivial (recall our convention that reflective subcategories are replete). Conversely, if (iii) holds, then given a diagram as in 4.3.4(*a*) we obtain a unique factorization $g': LB' \to A$ of f' through $\eta_{B'}$; but Lm is c_L-closed (since its domain and codomain are in $\mathcal{L}$),

and so on applying 4.3.3(i) to the square

$$\begin{tikzcd} B' \arrow[r, "\eta_{B'}"] \arrow[d, tail, "m"] & LB' \arrow[d, tail, "Lm"] \\ B \arrow[r, "\eta_B"] & LB \end{tikzcd}$$

we obtain a unique factorization of $\eta_{B'}$ through m, and hence a factorization of $f' = g'\eta_{B'}$ through m (the uniqueness of the latter follows from part (a)). □

It should be remarked that the equivalence of (i) and (iv) in 4.3.6(a) holds for any universal closure operation, whether or not it is induced by a cartesian reflector. For a direct proof of (iv) from (i), let $\overline{A} \rightarrowtail A \times A$ be the closure of the diagonal, and observe that the two projections $\overline{A} \rightrightarrows A$ are equalized by the dense monic $A \rightarrowtail \overline{A}$.

The next result is reminiscent of 2.6.9 (indeed, we shall see before long that 2.6.9 is a special case of it).

Lemma 4.3.7 *Let c be a universal closure operation on a cartesian category $\mathcal{E}$, and $m\colon B \rightarrowtail A$ a c-dense monomorphism in $\mathcal{E}$. Then pullback along m induces a bijection from (isomorphism classes of) closed subobjects of A to closed subobjects of B.*

Proof Let $B' \rightarrowtail B$ be a closed subobject. If there is any closed $A' \rightarrowtail A$ such that

$$\begin{tikzcd} B' \arrow[r, tail] \arrow[d, tail] & A' \arrow[d, tail] \\ B \arrow[r, tail, "m"] & A \end{tikzcd}$$

is a pullback, then $B' \rightarrowtail A'$ must be dense, and so by 4.3.3(ii) A' must be the closure of the composite monic $B' \rightarrowtail B \rightarrowtail A$. But if we define $A' \rightarrowtail A$ to be the closure of this composite, then $m^*(A')$ is the closure of $B' \rightarrowtail B$, and so is isomorphic to B'; that is, the square above is a pullback. □

Lemma 4.3.8 *Let c be a universal closure operation on a cartesian category $\mathcal{E}$, and let A be a c-sheaf. Then a subobject $m\colon A' \rightarrowtail A$ is c-closed iff its domain A' is also a c-sheaf.*

Proof Suppose m is closed, and suppose given a diagram

$$\begin{tikzcd} B' \arrow[r, "f'"] \arrow[d, tail, "n"] & A' \\ B & \end{tikzcd}$$

where n is dense. Since A is a sheaf, there is a unique $g: B \to A$ satisfying $gn = mf'$, and then 4.3.3(i) yields a unique factorization $f : B \to A'$ of g through m. So A' is a sheaf.

Conversely, suppose A' is a sheaf; let $\overline{m}: A'' \rightarrowtail A$ be the c-closure of m. Then $A' \rightarrowtail A''$ is dense, and so has a splitting $r: A'' \to A'$; and then the composite $mr: A'' \to A$ equals $\overline{m}$ since the two morphisms are equalized by the dense monomorphism $A' \rightarrowtail A''$. Since $\overline{m}$ is monic, it follows that r is a two-sided inverse for $A' \rightarrowtail A''$; so the latter is an isomorphism, and A' is closed in A. □

We have now assembled all the ingredients for the main result of this section.

Theorem 4.3.9 *Let $\mathcal{E}$ be a topos, and L a cartesian reflector on $\mathcal{E}$, corresponding to a reflective subcategory $\mathcal{L}$. Then $\mathcal{L}$ is a topos, and the inclusion $\mathcal{L} \to \mathcal{E}$ is the direct image of a geometric morphism, whose inverse image is (the factorization through $\mathcal{L}$ of) L.*

Proof $\mathcal{L}$ is cartesian since it is closed under limits in $\mathcal{E}$, and it is cartesian closed by 4.3.1. So it suffices to show that it has a subobject classifier. Let $J \rightarrowtail \Omega$ be the c_L-closure of the generic subobject $\top: 1 \rightarrowtail \Omega$, and $j: \Omega \to \Omega$ its classifying map. Then since c_L is stable under pullback it is easy to see that the entire closure operation is induced by j, in the sense that if $\phi: A \to \Omega$ classifies a subobject $A' \rightarrowtail A$, then the c_L-closure of A' is classified by the composite $j\phi$. In particular, if we define $\Omega_j \rightarrowtail \Omega$ to be the equalizer of j and 1_Ω, then Ω_j 'classifies closed subobjects'; that is, a morphism $\phi: A \to \Omega$ factors through Ω_j iff the subobject it classifies is c_L-closed. It now follows immediately from Lemma 4.3.7 that Ω_j is a c_L-sheaf (equivalently, by 4.3.6(b), an object of $\mathcal{L}$), and from 4.3.8 that it is a subobject classifier for $\mathcal{L}$. The fact that the indicated functors form a geometric morphism is immediate. □

Thus there is a bijection between (isomorphism classes of) cartesian reflectors on a topos $\mathcal{E}$ and (equivalence classes of) geometric inclusions, as defined in the previous section, with codomain $\mathcal{E}$. We do not yet know that these concepts also correspond bijectively to universal closure operations on $\mathcal{E}$, since we have not shown that all such closure operations arise from cartesian reflectors; but we shall do this in the next section. Once we have done so, then 4.3.5 and 4.3.9 together will provide a full proof of the result, only sketched in 2.1.10, that the category of sheaves on an arbitrary small site is a topos. In view of Example 4.2.12(c),

we commonly refer to the reflective subcategories of a topos $\mathcal{E}$ which correspond to cartesian reflectors as *subtoposes* of $\mathcal{E}$ (or *geometric subtoposes*, if we need to distinguish them from *logical subtoposes* which are full subcategories closed under finite limits and power objects, like $\mathbf{Set}_f$ in $\mathbf{Set}$; we have seen that a geometric subtopos is closed under finite limits and exponentials, but it cannot be closed under power objects (equivalently, contain Ω) unless it is the whole of $\mathcal{E}$, by the remark after 4.1.2).

Remark 4.3.10 It is instructive to re-examine the factorization of 4.2.10 from the point of view of cartesian reflectors and universal closure operations. Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism. Adopting the notation of 4.2.8, we see that the reflector $L = h_* h^* : \mathcal{E} \to \mathcal{E}$ sends an object A to the equalizer of

$$f_* f^* A \underset{\eta_{f_* f^* A}}{\overset{f_* f^* \eta_A}{\rightrightarrows}} f_* f^* f_* f^* A$$

(where η is the unit of $(f^* \dashv f_*)$), the unit map $A \to LA$ being the factorization of η_A through this equalizer. That is, the monad on $\mathcal{E}$ induced by $(h^* \dashv h_*)$ is the *associated idempotent monad* (in the sense of Fakir [347]) of the monad induced by $(f^* \dashv f_*)$. This idea can be used, in conjunction with 4.3.9, to provide an alternative construction of the surjection–inclusion factorization of an arbitrary geometric morphism, constructing the image as a subtopos of $\mathcal{E}$ rather than as a category of coalgebras over $\mathcal{F}$.

Now let $A' \rightarrowtail A$ be a monomorphism in $\mathcal{E}$. In the diagram

$$\begin{tikzcd} LA' \arrow[r, tail] \arrow[d, tail] & f_* f^* A' \arrow[r, shift left] \arrow[r, shift right] \arrow[d, tail] & f_* f^* f_* f^* A' \arrow[d, tail] \\ LA \arrow[r, tail] & f_* f^* A \arrow[r, shift left] \arrow[r, shift right] & f_* f^* f_* f^* A \end{tikzcd}$$

the two rows are equalizers, and since the vertical arrows are monic one readily deduces that the left-hand square is a pullback. Thus the universal closure operation c_L induced by L may be described directly in terms of f: specifically $c_L(A' \rightarrowtail A)$ is the pullback along $\eta_A : A \to f_* f^* A$ of $f_* f^* A' \rightarrowtail f_* f^* A$. Note that 4.2.6(v) provides a direct proof that this closure operation is trivial if f is a surjection.

It will frequently be useful, later on, to have a criterion for an arbitrary geometric morphism to factor through a given inclusion.

Proposition 4.3.11 *Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and L a cartesian reflector on $\mathcal{E}$. The following are equivalent:*

(i) *f factors (uniquely) through the inclusion $h: \mathcal{L} \to \mathcal{E}$ which corresponds to L under* 4.3.9.

(ii) *$f^*(m)$ is an isomorphism for every c_L-dense monomorphism m.*

(iii) *$f_*(B)$ is a c_L-sheaf for every object B of $\mathcal{F}$.*

Proof (i)$\Rightarrow$(ii): From the definition of c_L, it is clear that the functor $L = h_*h^*$ sends c_L-dense monomorphisms to isomorphisms. But h_* reflects isomorphisms, so h^* also has this property; hence so does f^* if there is a factorization $f^* \cong g^*h^*$.

(ii)$\Rightarrow$(iii) is immediate from the definition of a sheaf: given

$$\begin{tikzcd} A' \arrow[r, "g'"] \arrow[d, tail, "m"] & f_*(B) \\ A & \end{tikzcd}$$

with m c_L-dense, we transpose across the adjunction $(f^* \dashv f_*)$ to obtain

$$\begin{tikzcd} f^*(A') \arrow[r, "\overline{g'}"] \arrow[d, tail, "f^*(m)"] & B \\ f^*(A) & \end{tikzcd} .$$

But $f^*(m)$ is an isomorphism, so there is a unique $\overline{g}\colon f^*(A) \to B$ making this diagram commute; its transpose $g\colon A \to f_*(B)$ is the morphism we seek.

(iii)$\Rightarrow$(i): If (iii) holds, then by 4.3.6(b) there is a unique factorization $f_* = h_*g_*$. Moreover, the composite $g^* = f^*h_*$ is a cartesian left adjoint for g_*, since h_* is full and faithful. So we have a factorization $f \cong hg$ in $\mathfrak{Top}$. □

Corollary 4.3.12 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and let L and M be cartesian reflectors on $\mathcal{E}$ and $\mathcal{F}$ respectively, corresponding to subtoposes $\mathcal{L}$ and $\mathcal{M}$. The following are equivalent:*

(i) *There exists a (unique) geometric morphism $\mathcal{M} \to \mathcal{L}$ making the diagram*

$$\begin{tikzcd} \mathcal{M} \arrow[r] \arrow[d] & \mathcal{L} \arrow[d] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

commute up to isomorphism.

(ii) *f^* sends c_L-dense monomorphisms to c_M-dense monomorphisms.*

(iii) *f_* sends c_M-sheaves to c_L-sheaves.*

Proof These are just the three equivalent conditions of 4.3.11 applied to the composite $\mathcal{M} \to \mathcal{F} \to \mathcal{E}$. (Note that the monomorphisms in $\mathcal{F}$ which are sent to isomorphisms in $\mathcal{M}$ by the inverse image functor are precisely the $c_{\mathcal{M}}$-dense ones.) □

Suggestions for further reading: Bell & Gebellato [98], Borceux [141, 144], Borceux & Kelly [151], Borceux & Van Den Bossche [164], Johnstone [538, 539], Kelly & Lawvere [586], Wyler [1248].

A4.4 Local operators

In this section we study universal closure operations on a topos; our primary aim is to show that they are all induced by cartesian reflectors, and so correspond to subtoposes via 4.3.9. However, for the sake of completeness we shall also include a number of results about universal closure operators on quasitoposes.

In the course of proving 4.3.9, we saw that a universal closure operation c on a topos is entirely determined by a particular endomorphism of Ω, namely the classifying map of the closure of the generic subobject $\top$. It is convenient to have a name for the endomorphisms of Ω which arise in this way; we shall call them local operators. (The old name for them is *(Lawvere–Tierney) topologies*, but this name can lead to confusion since they have little to do with topologies in the classical sense.)

Definition 4.4.1 Let $\mathcal{E}$ be a topos, with subobject classifier Ω. A *local operator* on $\mathcal{E}$ (or on Ω) is a morphism $j\colon \Omega \to \Omega$ such that the diagrams

$$\begin{tikzcd} 1 \arrow[r, "\top"] \arrow[dr, "\top"'] & \Omega \arrow[d, "j"] \\ & \Omega \end{tikzcd} \quad , \quad \begin{tikzcd} \Omega \arrow[r, "j"] \arrow[dr, "j"'] & \Omega \arrow[d, "j"] \\ & \Omega \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} \Omega \times \Omega \arrow[r, "\wedge"] \arrow[d, "j \times j"] & \Omega \arrow[d, "j"] \\ \Omega \times \Omega \arrow[r, "\wedge"] & \Omega \end{tikzcd}$$

commute (where $\wedge$ is the binary meet operation on Ω defined in 1.6.3).

Lemma 4.4.2 *For any topos $\mathcal{E}$, there is a bijection between universal closure operations on $\mathcal{E}$ and local operators on $\mathcal{E}$.*

Proof Given $j\colon \Omega \to \Omega$, let c_j denote the (pullback-stable) unary operation on subobjects in $\mathcal{E}$ induced by composing classifying maps with j. We have to show that c_j is a universal closure operation iff j is a local operator. But the first two diagrams of 4.4.1 are precisely equivalent to saying that $A' \le c_j(A') \cong c_j c_j(A')$ for any subobject $A' \rightarrowtail A$ in $\mathcal{E}$; the third diagram says that c_j commutes with binary intersection of subobjects, which in particular implies that it is order-preserving. Conversely, we saw in 4.3.3(iii) that any universal closure operation

commutes with binary intersections, so the endomorphism of Ω which induces it must be a local operator. □

Given a local operator j, we adopt as standard the notation already introduced in the proof of 4.3.9; that is, we write $J \rightarrowtail \Omega$ for the subobject classified by j, and $\Omega_j \rightarrowtail \Omega$ for the equalizer of j and 1_Ω. Note that since j is idempotent, we may equivalently describe Ω_j as the image of j; in particular, it is a retract of Ω, and therefore injective in $\mathcal{E}$ by 2.2.6. As in 4.3.9, we see that Ω_j 'classifies closed subobjects' for the corresponding universal closure operation; and J similarly classifies dense subobjects – that is, a morphism $\phi\colon A \to \Omega$ factors through J iff $j\phi$ factors through $\top$, iff the subobject of A classified by ϕ is c_j-dense.

We may also talk about local operators on a quasitopos $\mathcal{E}$ (defined exactly as in 4.4.1); but for a quasitopos the relationship between universal closure operations and local operators is slightly more complicated. We shall say that a universal closure operation on a quasitopos is *proper* if the closure of a cocover is always a cocover; we note that the closure operation c_L induced by a cartesian reflector L always has this property, since cocovers coincide with regular monomorphisms and L preserves equalizers. If c is a proper universal closure operation, then its action on cocovers is determined as in 4.4.2 by a local operator $j\colon \Omega \to \Omega$. However, we cannot hope to recover the whole of c from j: for example, let c_1 be the closure operation such that, for any $A' \rightarrowtail A$, $A' \rightarrowtail c_1(A') \rightarrowtail A$ is the unique factorization of $A' \rightarrowtail A$ as a (monic) epimorphism followed by a cocover. (This factorization is stable under pullback, since epimorphisms and cocovers both are. Moreover, the c_1-sheaves are exactly the coarse objects of $\mathcal{E}$, as defined in Section A2.6; so by 2.6.12 and 4.3.6(*b*) this closure operation corresponds to a reflector on $\mathcal{E}$. We shall see shortly that this reflector is in fact cartesian.) Then c_1 sends each cocover to itself, and so the local operator which it induces is the identity; but c_1 is not the identity operation on subobjects, unless $\mathcal{E}$ is a topos.

We shall say that c is *strict* if the closure of every subobject is a cocover. Given a proper universal closure operation c, we can construct a strict closure operation $\bar{c}$ having the same effect on cocovers as c, by defining $\bar{c}(A') = c(c_1(A'))$ where c_1 is as above. There is a bijection between local operators $j\colon \Omega \to \Omega$ and strict universal closure operations on $\mathcal{E}$: given j, we define $c_j(A' \rightarrowtail A)$ to be the cocover classified by $j\phi$, where ϕ is the classifying map of $c_1(A' \rightarrowtail A)$. We further note that the passage from c to $\bar{c}$ above does not affect the notion of separated object: since every $\bar{c}$-dense monomorphism factors as an epimorphism followed by a c-dense cocover, we see that every c-separated object is $\bar{c}$-separated (and the converse is trivial). However, in general $\bar{c}$ will have fewer sheaves than c.

Given a universal closure operation c on a cartesian category $\mathcal{E}$, we shall write $\mathbf{sep}_c(\mathcal{E})$ and $\mathbf{sh}_c(\mathcal{E})$ for the full subcategories of c-separated objects and c-sheaves in $\mathcal{E}$.

Lemma 4.4.3 *Let c be a universal closure operation on a cartesian category $\mathcal{E}$. Then*

(i) *The subcategories* $\mathbf{sep}_c(\mathcal{E})$ *and* $\mathbf{sh}_c(\mathcal{E})$ *are closed under finite limits in* $\mathcal{E}$.
(ii) *If $\mathcal{E}$ is cartesian closed, then* $\mathbf{sep}_c(\mathcal{E})$ *and* $\mathbf{sh}_c(\mathcal{E})$ *are exponential ideals in* $\mathcal{E}$.

Proof (i) This is immediate from the fact that the conditions for A to be a sheaf or a separated object involve only morphisms with codomain A.

(ii) Suppose A is a c-sheaf and B is any object of $\mathcal{E}$. For any c-dense $m: C' \rightarrowtail C$, the morphism $m \times 1_B : C' \times B \rightarrowtail C \times B$ is also c-dense (since it is the pullback of m along the product projection), and so any morphism $C' \times B \to A$ extends uniquely to a morphism $C \times B \to A$. But this is equivalent to saying that A^B is a sheaf. The argument for separated objects is similar. □

Now suppose c is a proper universal closure operation on a quasitopos $\mathcal{E}$. Write $j: \Omega \to \Omega$ for the local operator induced by c (i.e. by the action of c on cocovers), and $\Omega_j \rightarrowtail \Omega$ for the equalizer of j and 1_Ω. As before, we see that morphisms $A \to \Omega_j$ correspond to monomorphisms $A' \rightarrowtail A$ which are cocovers and c-closed, i.e. to $\bar{c}$-closed subobjects of A. So Lemma 4.3.7 implies that Ω_j is a $\bar{c}$-sheaf, and in particular a c-sheaf. We shall write $P_j(A)$ for the exponential Ω_j^A; by 4.4.3(ii) this is also a $\bar{c}$-sheaf.

Proposition 4.4.4 *Let c be a proper universal closure operation on a quasitopos $\mathcal{E}$. Then the subcategories* $\mathbf{sep}_c(\mathcal{E})$ *and* $\mathbf{sh}_c(\mathcal{E})$ *are both reflective in* $\mathcal{E}$.

Proof By definition, they are both full and replete, so it suffices to construct the appropriate left adjoints. Let A be an arbitrary object of $\mathcal{E}$; write $\overline{A} \rightarrowtail A \times A$ for the $\bar{c}$-closure of the diagonal $\Delta: A \rightarrowtail A \times A$ (which in fact coincides with its c-closure, since Δ is regular monic), and let $d: A \to P_j(A)$ be the 'name' of this $\bar{c}$-closed relation. We note in passing that $\overline{A}$ is an equivalence relation on A, as defined in 1.3.6; this follows easily from the pullback-stability of the closure operation. We write $MA \rightarrowtail P_j(A)$ for the image of d (in the sense of the cover–monic factorization, not the epic–cocover one!), and $LA \rightarrowtail P_j(A)$ for the c-closure of MA. Since $P_j(A)$ is a c-sheaf, it follows from 4.3.8 that LA is a c-sheaf; and MA is c-separated, since it is easily seen that any subobject of a separated object is separated.

By (the quasitopos version of) Proposition 2.4.1, the kernel-pair of d (and hence of the regular epimorphism $A \twoheadrightarrow MA$) is $\overline{A} \rightrightarrows A$; so a morphism $f: A \to B$ factors (uniquely) through $A \twoheadrightarrow MA$ iff it coequalizes these two projections. But this is always the case if B is c-separated, since the two projections are equalized by the c-dense monic $A \rightarrowtail \overline{A}$. Thus MA is the reflection of A in $\mathbf{sep}_c(\mathcal{E})$. If further B is a c-sheaf, then any morphism $MA \to B$ can be uniquely extended along the c-dense monic $MA \rightarrowtail LA$, from which it follows that LA is the reflection of A in $\mathbf{sh}_c(\mathcal{E})$. □

The reflection LA is commonly called the *associated sheaf* (or sometimes the *sheafification*) of the object A.

Theorem 4.4.5 *Let c be a proper universal closure operation on a quasitopos $\mathcal{E}$. Then the categories $\mathbf{sep}_c(\mathcal{E})$ and $\mathbf{sh}_c(\mathcal{E})$ are both quasitoposes. Further, if c is strict, then $\mathbf{sh}_c(\mathcal{E})$ is a topos.*

Proof By 4.4.3(i) both subcategories inherit finite limits from $\mathcal{E}$, and by 4.4.4 they are both cocartesian. By 4.4.3(ii), they are both cartesian closed; but we have to show that they are locally cartesian closed. To this end, let B be an object of $\mathbf{sep}_c(\mathcal{E})$, and define a universal closure operation c_B on $\mathcal{E}/B$ by setting

$$c_B((A' \to B) \rightarrowtail (A \to B)) = ((c(A') \to B) \rightarrowtail (A \to B))$$

where $c(A')$ denotes the c-closure of A' in A. It is straightforward to verify that an object $(A \to B)$ of $\mathcal{E}/B$ is c_B-separated iff A is c-separated in $\mathcal{E}$; so we have an isomorphism

$$\mathbf{sep}_c(\mathcal{E})/B \cong \mathbf{sep}_{c_B}(\mathcal{E}/B),$$

from which it follows that $\mathbf{sep}_c(\mathcal{E})/B$ is cartesian closed. The argument for sheaves is similar.

Finally, we must determine the cocovers in $\mathbf{sep}_c(\mathcal{E})$ and $\mathbf{sh}_c(\mathcal{E})$; we shall show that they are exactly the $\bar{c}$-closed monomorphisms, so that Ω_j will serve as a weak subobject classifier for both categories. Moreover, in the case when c is strict, so that all c-closed monomorphisms are cocovers and hence $\bar{c}$-closed, Ω_j is actually a subobject classifier for $\mathbf{sh}_c(\mathcal{E})$, since all monomorphisms in this category are c-closed by 4.3.8. One direction is easy: any c-dense monomorphism is epic in $\mathbf{sep}_c(\mathcal{E})$, by the very definition of a separated object, and since any $\bar{c}$-dense monomorphism factors as an epimorphism in $\mathcal{E}$ (which is clearly still epic in $\mathbf{sep}_c(\mathcal{E})$) followed by a c-dense monomorphism, it follows that a cocover in $\mathbf{sep}_c(\mathcal{E})$ must be $\bar{c}$-closed. In $\mathbf{sh}_c(\mathcal{E})$ the argument is even easier: here we know already that all monomorphisms are c-closed, so it suffices to prove that cocovers in $\mathbf{sh}_c(\mathcal{E})$ must be cocovers in $\mathcal{E}$ – which is immediate. For the converse, it suffices to observe that every $\bar{c}$-closed monomorphism is a pullback (in $\mathcal{E}$, and hence in $\mathbf{sep}_c(\mathcal{E})$ or $\mathbf{sh}_c(\mathcal{E})$ as appropriate) of the generic $\bar{c}$-closed monomorphism $\top_j : 1 \rightarrowtail \Omega_j$ (that is, the unique factorization of $\top$ through Ω_j), which is split and hence regular monic in any full subcategory of $\mathcal{E}$ in which it lives. □

We note that Theorem 4.4.5 includes as particular cases a number of results which we have seen before. For example, if we take c to be the particular (strict) closure operation c_1 mentioned earlier, it yields the result of 2.6.12 that $\mathbf{Cs}(\mathcal{E})$ is a topos for every quasitopos $\mathcal{E}$. When $\mathcal{E}$ itself is a topos (so that every universal closure operation on $\mathcal{E}$ is strict), it reduces to the result of 4.3.9 that every reflective subcategory of $\mathcal{E}$ with cartesian reflector is a topos. (Note: when $\mathcal{E}$ is a topos, it is customary to designate subcategories of separated objects or sheaves by the name of the corresponding local operator rather than that of the universal

closure operation: that is, we write $\mathbf{sep}_j(\mathcal{E})$ and $\mathbf{sh}_j(\mathcal{E})$ rather than $\mathbf{sep}_c(\mathcal{E})$ and $\mathbf{sh}_c(\mathcal{E})$.) And when $\mathcal{E}$ is a topos of the form $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, then (in conjunction with 4.3.5) 4.4.5 yields the results of 2.6.4(d) and 2.1.10 that $\mathbf{Sep}(\mathcal{C}, T)$ is a quasitopos and $\mathbf{Sh}(\mathcal{C}, T)$ is a topos for any small site $(\mathcal{C}, T)$.

Next, we investigate the limit-preservation properties of the two reflectors L and M constructed in the proof of 4.4.4. By 4.3.1 and 4.4.3(ii), we already know that they both preserve finite products.

Lemma 4.4.6 *The functor L is cartesian.*

Proof It suffices to show that it preserves equalizers. Let

$$A \overset{e}{\rightarrowtail} B \underset{g}{\overset{f}{\rightrightarrows}} C$$

be an equalizer diagram in $\mathcal{E}$, and let $d\colon E \rightarrowtail LB$ be the equalizer of Lf and Lg. Since L is a functor, we have a unique factorization $k\colon LA \to E$ of Le through d; we must show that k is an isomorphism.

First we show that k is monic. Let $(u, v)\colon X \rightrightarrows LA$ be a pair of morphisms with $ku = kv$. On forming the pullback

$$\begin{tikzcd} P \arrow[r, "r"] \arrow[d, "{(p,q)}"] & X \arrow[d, "{(u,v)}"] \\ A \times A \arrow[r, "\eta_A \times \eta_A"] & LA \times LA \end{tikzcd}$$

we obtain a pair (p, q) with

$$\eta_B ep = (Le)\eta_A p = dkur = dkvr = (Le)\eta_A q = \eta_B eq\,.$$

So the pair $(ep, eq)\colon P \rightrightarrows B$ factors through the kernel-pair $\overline{B} \rightrightarrows B$ of η_B; but by the construction of 4.4.4 the diagonal map $B \rightarrowtail \overline{B}$ is dense, and so on pulling it back we get a dense $P' \rightarrowtail P$ equalizing ep and eq (and hence equalizing p and q, since e is monic). Since LA is separated, it follows that $\eta_A p = \eta_A q$, i.e. $ur = vr$; but the image of r is dense in X, because the image of $\eta_A \times \eta_A$ (namely $MA \times MA$) is dense in $LA \times LA$, and hence $u = v$, as required.

Now LA and E are both sheaves, so by 4.3.8 the monomorphism $k\colon LA \rightarrowtail E$ is closed. We shall show that it is also dense, and hence an isomorphism. Form the pullback

$$\begin{tikzcd} Q \arrow[r, "h"] \arrow[d, "i"] & B \arrow[d, "\eta_B"] \\ E \arrow[r, "d"] & LB \end{tikzcd} \quad ;$$

then we have

$$\eta_C fh = (Lf)\eta_B h = (Lf)di = (Lg)di = (Lg)\eta_B h = \eta_C gh,$$

so the pair $(fh, gh)\colon Q \rightrightarrows C$ factors through the kernel-pair $\overline{C} \rightrightarrows C$ of η_C. As in the previous paragraph, this implies that there is a dense $m\colon Q' \rightarrowtail Q$ equalizing fh and gh, so that the composite $hm\colon Q' \to B$ factors through e (say by $l'\colon Q' \to A$). Now since LA is a sheaf, the composite $\eta_A l'\colon Q' \to LA$ extends uniquely to a morphism $l\colon Q \to LA$; and since E is separated, the diagram

$$\begin{tikzcd} Q \arrow[r, "l"] \arrow[dr, "i"'] & LA \arrow[d, tail, "k"] \\ & E \end{tikzcd}$$

commutes, for we have

$$dim = \eta_B hm = \eta_B el' = (Le)\eta_A l' = (Le)lm = dklm$$

and d is monic and m is dense. So, in $\mathrm{Sub}(E)$, k contains the image of i; but the latter is dense, since it is the pullback along d of the image of η_B. □

Remark 4.4.7 For toposes, there is a much shorter (but more 'high-technology') proof of 4.4.6 using properties of injective objects. First we note that, by 2.2.6, the injective objects of $\mathbf{sh}_j(\mathcal{E})$ are exactly the retracts of power objects Ω_j^A, $A \in \mathrm{ob}\,\mathbf{sh}_j(\mathcal{E})$; but since Ω_j is a retract of Ω in $\mathcal{E}$, all such objects are in fact injective in $\mathcal{E}$. Again by 2.2.6, we note that a morphism $m\colon A \to B$ in a topos is monic iff every morphism from A to an injective object factors through m; from this, and the fact that the inclusion $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$ preserves injectives, it is easy to see that L preserves monomorphisms.

We now claim that L preserves binary intersections of subobjects (and hence in particular equalizers of coreflexive pairs, cf. 1.2.10); given that it preserves finite products by 4.3.1 and 4.4.3(ii), this is enough to ensure that it preserves all finite limits. Suppose given subobjects $A_1 \rightarrowtail A$ and $A_2 \rightarrowtail A$ in $\mathcal{E}$; then by 1.4.3 we have a pushout diagram

$$\begin{tikzcd} A_1 \cap A_2 \arrow[r, tail] \arrow[d, tail] & A_1 \arrow[d, tail] \\ A_2 \arrow[r, tail] & A_1 \cup A_2 \end{tikzcd}$$

in $\mathcal{E}$, and hence the square

$$\begin{tikzcd} L(A_1 \cap A_2) \arrow[r, tail] \arrow[d, tail] & LA_1 \arrow[d, tail] \\ LA_2 \arrow[r, tail] & L(A_1 \cup A_2) \end{tikzcd}$$

is a pushout in $\mathbf{sh}_j(\mathcal{E})$ since L is a left adjoint. By 2.4.3 (or by 1.4.8), this square is also a pullback; but since $L(A_1 \cup A_2) \to LA$ is monic, the square

$$\begin{tikzcd} L(A_1 \cap A_2) \arrow[r, tail] \arrow[d, tail] & LA_1 \arrow[d, tail] \\ LA_2 \arrow[r, tail] & LA \end{tikzcd}$$

is also a pullback, i.e. $L(A_1 \cap A_2)$ is the intersection of LA_1 and LA_2 in $\mathrm{Sub}(LA)$.

The first part of the above argument works for quasitoposes as well as for toposes: note that powers of Ω in a quasitopos are injective for arbitrary monomorphisms (not just for cocovers), because they are coarse objects by 2.6.9. It can therefore be used to give a proof that the reflector $M\colon \mathcal{E} \to \mathbf{sep}_c(\mathcal{E})$ preserves monomorphisms, for any proper universal closure operation c on a quasitopos $\mathcal{E}$; however, the same result could alternatively be deduced from the fact that M is a subfunctor of L and L preserves monomorphisms. On the other hand, the second part of the argument above definitely requires the reflective subcategory of $\mathcal{E}$ to be a (pre-)topos, since it requires the result of 2.4.3 to hold for pushouts of arbitrary monomorphisms, and not just for cocovers (cf. 2.6.2).

Theorem 4.4.8 *For a quasitopos $\mathcal{E}$, there is a bijection between reflective subcategories of $\mathcal{E}$ with cartesian reflector, and proper universal closure operations on $\mathcal{E}$. In particular, if $\mathcal{E}$ is a topos, there is a bijection between subtoposes of $\mathcal{E}$ and local operators on $\mathcal{E}$.*

Proof We saw in 4.3.2 how every cartesian reflector gives rise to a universal closure operation (and we observed before 4.4.3 that this closure operation is proper); conversely, the combination of 4.4.4 and 4.4.6 shows how every proper universal closure operation gives rise to a cartesian reflector. If we start from a cartesian reflector and apply these two constructions in succession, we arrive back at the same reflective subcategory, by 4.3.6(b). So it remains to show that, if we start from a proper universal closure operation c and perform the two constructions in the other order, we recover c; that is, for any monomorphism $m\colon A' \rightarrowtail A$ in $\mathcal{E}$, the pullback $A'' \rightarrowtail A$ of Lm along η_A is isomorphic to $c(A')$ in $\mathrm{Sub}(A)$. But $A'' \rightarrowtail A$ is closed, since Lm is closed by 4.3.8, so by 4.3.3(ii)

it suffices to show that $A' \rightarrowtail A''$ is dense. But the latter monomorphism is mapped by L to an isomorphism; so the induced morphism $MA' \to MA''$ is dense monic, since the composite $MA' \to MA'' \to LA'' \cong LA'$ is dense monic by the construction of 4.4.4. Now there is a pullback square

$$\begin{tikzcd} A' \arrow[r, tail] \arrow[d, tail] & A'' \arrow[d, tail] \\ A' \times_{MA''} A'' \arrow[r, tail] & A'' \times_{MA''} A'' \end{tikzcd}$$

in which the right vertical arrow (the diagonal) is dense monic by the construction of 4.4.4, so the left vertical arrow is dense monic. Moreover, the projection $A' \times_{MA''} A'' \to A''$ factors as a cover followed by a dense monomorphism, since it is a pullback of the composite $A' \twoheadrightarrow MA' \rightarrowtail MA''$; and a monomorphism which factors as a dense monic followed by a cover must be dense, since pullback along a cover, being conservative, reflects as well as preserving dense monomorphisms. Thus we have expressed the monomorphism $A' \rightarrowtail A''$ as a composite of two dense monomorphisms; so it is dense, as claimed.

The second assertion of the theorem follows immediately from the first and 4.4.2. □

Thus we see that a proper universal closure operation c on a quasitopos $\mathcal{E}$ is uniquely determined by the subcategory $\mathbf{sh}_c(\mathcal{E})$. It is not uniquely determined by $\mathbf{sep}_c(\mathcal{E})$, since we have already observed that $\mathbf{sep}_c(\mathcal{E}) = \mathbf{sep}_{\bar{c}}(\mathcal{E})$ for any c; but if we add the requirement that c should be strict, then $\mathbf{sep}_c(\mathcal{E})$ does determine it, since it is easily seen that in this case we have $\mathbf{sh}_c(\mathcal{E}) = \mathbf{Cs}(\mathbf{sep}_c(\mathcal{E}))$. In particular, if $\mathcal{E}$ is a topos, then $\mathbf{sep}_j(\mathcal{E})$ determines the local operator j.

We saw in 4.4.7 that the separated reflector M preserves monomorphisms, but it is not cartesian in general. In fact if M is cartesian, we see from 4.4.8 that $\mathbf{sep}_c(\mathcal{E})$ must coincide with $\mathbf{sh}_{c'}(\mathcal{E})$ for some c'; moreover, since arbitrary subobjects of c-separated objects are separated, it follows from 4.3.6(a) that every separated object for c' must be a sheaf for it (and so, if $\mathcal{E}$ is a topos, c' must coincide with c). Nontrivial examples of universal closure operations c with $\mathbf{sep}_c(\mathcal{E}) = \mathbf{sh}_c(\mathcal{E})$ can exist, even when $\mathcal{E}$ is a topos; but they are not very common. Here is a simple one:

Example 4.4.9 Let M be the two-element monoid $\{1, e\}$ with $e^2 = e$, and let $\mathcal{E}$ be the topos of M-sets (that is, sets equipped with an idempotent endomorphism). Let $f_* : \mathbf{Set} \to \mathcal{E}$ be the functor which equips a set with its identity endomorphism, and $f^* : \mathcal{E} \to \mathbf{Set}$ the functor which sends (A, e) to the set of fixed points of e. It is well known that f^* and f_* are adjoint to each other both ways round, and so they form a geometric morphism in either direction between $\mathcal{E}$ and $\mathbf{Set}$. In the direction indicated by the notation, this morphism is an inclusion, since f_* is full and faithful; but if we use it to identify $\mathbf{Set}$ with a subtopos

of $\mathcal{E}$, we see that the latter is closed under arbitrary subobjects in $\mathcal{E}$, so that every separated object for the corresponding local operator must be a sheaf.

We conclude this section by briefly mentioning an alternative construction of the reflector $\mathcal{E} \to \mathbf{sh}_j(\mathcal{E})$ for a local operator j on a topos $\mathcal{E}$, which is described in greater detail in [501]. As previously, let $J \rightarrowtail \Omega$ denote the subobject classified by j, and let $d\colon 1 \rightarrowtail J$ be the factorization of $\top\colon 1 \rightarrowtail \Omega$ through J (the generic dense subobject). Let $\delta\colon \hat{A} \to J$ denote $\Pi_d(A)$; equivalently, it is the pullback of $\tau\colon \tilde{A} \to \Omega$ along $J \rightarrowtail \Omega$ (cf. 2.4.7). Just as in the proof of 2.4.7, we may show that $\hat{A}$ classifies partial maps into A with dense domain. The alternative construction involves making $\hat{A} \to J$ into a (contravariant) internal diagram on the internal poset $\mathbf{J} = (J_1 \rightrightarrows J)$ (where J_1 is the intersection of $J \times J$ with the ordering $\Omega_1 \rightarrowtail \Omega \times \Omega$ defined in the proof of 1.6.3), and forming its colimit A^+; it can then be shown that, for any object A, A^{++} is the reflection of A in $\mathbf{sh}_j(\mathcal{E})$. Since we have not yet developed the theory of internal categories and diagrams, we shall not give any further details of this construction; but we note that its 'external' version, for sheaves on a site, may be found in C2.2.6. For future reference, we also note the following characterization of sheaves and separated objects, which is used in the proof:

Lemma 4.4.10 *Let j be a local operator on a topos $\mathcal{E}$, A an object of $\mathcal{E}$. Let $\psi_A\colon A \times J \to \hat{A}$ classify the partial map*

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, tail, "1 \times d"] & A \\ A \times J & \end{tikzcd}$$

Then A is separated (resp. a sheaf) iff ψ_A is monic (resp. an isomorphism).

Proof For any object B, morphisms $B \to A \times J$ and $B \to \hat{A}$ correspond respectively to diagrams

$$\begin{tikzcd} B' \arrow[d, tail, "m"] & A \\ B \arrow[ur, "f"'] & \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} B' \arrow[r, "g"] \arrow[d, tail, "m"] & A \\ B & \end{tikzcd}$$

with m j-dense; and composition with ψ_A corresponds to the operation which sends (m, f) to (m, fm). So this is immediate from the definitions. □

Suggestions for further reading: Betti & Carboni [115], Bucalo & Rosolini [178], Dupont & Loiseau [317], Freyd [371], Goldblatt [412], Johnstone [501], Veit [1197].

A4.5 Examples of local operators

In this section we discuss some particular classes of local operators on a topos, and the subtoposes which correspond to them; we shall also give some results on the structure of the ordered set of all local operators on a topos. Much of this discussion, too, could be generalized to quasitoposes; but we shall not go into the details here.

To begin with, let U be a subterminal object in a topos $\mathcal{E}$. By 4.2.12(a), the geometric morphism $\mathcal{E}/U \to \mathcal{E}/1 \cong \mathcal{E}$ induced by $U \rightarrowtail 1$ is an inclusion, so we may identify the slice category $\mathcal{E}/U$ with a subtopos of $\mathcal{E}$. Note that the identification must be made, not via the forgetful functor Σ_U, but via the direct image Π_U of the geometric morphism. (However, since $U^U \cong 1$, as follows easily from 1.5.10, the description of Π_U given in 1.5.2(i) may be simplified in this case; $\Pi_U(f\colon A \to U)$ is simply the exponential A^U.) In the case where $\mathcal{E} = \mathbf{Sh}(X)$, and U is (identified with) an open subset of X, then $\mathcal{E}/U$ may be identified with $\mathbf{Sh}(U)$ (and the geometric morphism above is identified with that induced by the inclusion map $U \to X$); for this reason, subtoposes of this form are called *open subtoposes*.

The reflector on $\mathcal{E}$ corresponding to this subtopos is simply the composite $\Pi_U U^* \cong (-)^U$, and the unit map $A \to A^U$ is the transpose of the projection map $A \times U \to A$. It is thus easy to verify that the corresponding universal closure operation sends a subobject $A' \rightarrowtail A$ to the implication $(A \times U) \Rightarrow A'$ in the Heyting algebra $\mathrm{Sub}(A)$; equivalently, the local operator corresponding to the subtopos is the composite

$$\Omega \cong 1 \times \Omega \xrightarrow{u \times 1} \Omega \times \Omega \xrightarrow{\Rightarrow} \Omega,$$

where $u\colon 1 \to \Omega$ is the classifying map of $U \rightarrowtail 1$. It is also easy to verify directly from the axioms for a Heyting algebra that this composite satisfies the conditions of 4.4.1; we shall denote it by $o(U)$.

Open subtoposes may be characterized by a number of equivalent conditions; the following proposition should be compared with 2.3.8.

Proposition 4.5.1 *Let j be a local operator on a topos $\mathcal{E}$, and let $\mathcal{L}$ denote the topos of j-sheaves. The following conditions are equivalent:*

(i) *$\mathcal{L}$ is an open subtopos of $\mathcal{E}$.*

(ii) *The reflector $L\colon \mathcal{E} \to \mathcal{L}$ is a logical functor.*

(iii) *The reflector $L\colon \mathcal{E} \to \mathcal{L}$ is a cartesian closed functor.*

(iv) *If* $m\colon A' \rightarrowtail A$ *is* j*-dense, so is* $m^B\colon A'^B \rightarrowtail A^B$ *for any* B.

(v) *If* $d\colon 1 \rightarrowtail J$ *denotes the unique factorization of* $\top\colon 1 \rightarrowtail \Omega$ *through the subobject classified by* j, *then* $\Pi_J(d)$ *is a* j*-dense subobject of* 1.

(vi) *There exists a morphism* $u\colon 1 \to J$ *such that the morphism*

$$J \cong 1 \times J \xrightarrow{u \times 1} J \times J$$

factors through the order-relation $J_1 \rightarrowtail J \times J$ *(i.e. the restriction of the order-relation on* Ω*, as defined in* 1.6.3*).*

Proof (i) $\Rightarrow$ (ii) is immediate from 2.3.2, and (ii) $\Rightarrow$ (iii) follows from 2.3.7(iii). For (iii) $\Rightarrow$ (iv), recall that the j-dense monomorphisms are precisely those which are mapped by L to isomorphisms. (iv) implies that $d^J\colon 1^J \rightarrowtail J^J$ is j-dense, since d is dense by the definition of J; hence $\Pi_J(d) \rightarrowtail 1$ is j-dense by the construction of Π_J given in 1.5.2, i.e. (v) holds. For (v) $\Rightarrow$ (vi), let u be (the factorization through J of) the classifying map of $\Pi_J(d) \rightarrowtail 1$; since the counit of $(J^* \dashv \Pi_J)$ yields an inequality $J^*\Pi_J(d) \leq d$ in $\mathrm{Sub}(J)$, we deduce that the displayed morphism factors through the order-relation on J. Finally, if (vi) holds, let $U \rightarrowtail 1$ be the subobject classified by u; then we see that, for any object A of $\mathcal{E}$, $A \times U$ is the unique smallest j-dense subobject of A. So an object B is a j-sheaf iff, for any A, each morphism $A \times U \to B$ extends uniquely to a morphism $A \to B$; but this is equivalent to saying that every morphism $A \to B^U$ factors through the transpose $B \to B^U$ of the projection, and hence to saying that the latter is an isomorphism. So the j-sheaves are exactly the objects of the open subtopos corresponding to U. $\square$

Although, as we saw in the course of the above proof, an open local operator has the property that every object of the topos has a smallest dense subobject, the latter condition does not suffice to characterize open subtoposes; we need the extra information that the smallest dense subobject of A is of the form $A \times U$ for a fixed subterminal object U.

Example 4.5.2 Let $\mathcal{C}$ be a small category, and $\mathcal{D}$ a full subcategory of $\mathcal{C}$. Then the geometric morphism $[\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ induced by the inclusion $\mathcal{D} \to \mathcal{C}$ is an inclusion by 4.2.12(*b*); so it corresponds to a local operator on $[\mathcal{C}, \mathbf{Set}]$. It is not hard to see that a monomorphism $F' \rightarrowtail F$ in $[\mathcal{C}, \mathbf{Set}]$ is dense for this local operator iff $F'(b) \rightarrowtail F(b)$ is bijective for all objects b of $\mathcal{D}$; hence any F has a smallest dense subobject F', given by

$$F'(a) = \{x \in F(a) \mid \exists f : b \to a \text{ with } b \in \mathrm{ob}\,\mathcal{D} \text{ and } y \in F(b) \text{ with } F(f)(y) = x\}\ .$$

But the subterminal objects in $[\mathcal{C}, \mathbf{Set}]$ correspond to full subcategories of $\mathcal{C}$ which are *cosieves*, i.e. have the property that $b \in \mathrm{ob}\,\mathcal{D}$ and $f : b \to a$ in $\mathcal{C}$ imply $a \in \mathrm{ob}\,\mathcal{D}$. So if $\mathcal{D}$ is a full subcategory which is not a cosieve, we obtain an example of a local operator which is not open, but for which every object has a smallest dense subobject.

The next class of subtoposes we wish to consider is also indexed by the subterminal objects of a given topos; they play the rôle of the 'closed complements' of the open subtoposes we have just investigated, and we naturally call them *closed subtoposes*. Given $U \rightarrowtail 1$ in a topos $\mathcal{E}$, with classifying map $u\colon 1 \to \Omega$, consider the composite

$$\Omega \cong 1 \times \Omega \xrightarrow{u \times 1} \Omega \times \Omega \xrightarrow{\vee} \Omega$$

where $\vee$ is the binary join map on Ω, as defined in 1.6.3(ii). Since a Heyting algebra is distributive as a lattice, this composite satisfies the third commutative diagram of 4.4.1; the fact that it satisfies the first two is immediate, and so it is a local operator, which we shall denote by $\mathrm{c}(U)$. We shall see below that the local operators on a topos form a distributive lattice (in fact a Heyting algebra), and that the open and closed local operators corresponding to a given subterminal object are indeed complementary to each other in this lattice. For the present, we note that if U is a complemented subobject of 1, with complement V say, then $c(U)$ coincides with $o(V)$, by an easy calculation in the Heyting algebra Ω; thus, for a Boolean topos, the notions of open subtopos and closed subtopos coincide. (Again, we shall see later that if $\mathcal{E}$ is a Boolean topos then its only subtoposes are open ones.)

Proposition 4.5.3 *Let U be a subterminal object in a topos. Then*

(i) *A monomorphism $m\colon A' \rightarrowtail A$ is $c(U)$-dense iff A is the union of the subobjects A' and $A \times U$; and it is $c(U)$-closed iff it contains $A \times U$.*
(ii) *A is a $c(U)$-sheaf iff $A \times U \cong U$.*
(iii) *The reflection of an arbitrary object A in the subtopos of* $\mathrm{c}(U)$*-sheaves is the pushout of*

$$\begin{tikzcd} A \times U \arrow[r, "\pi_1"] \arrow[d, "\pi_2"] & A \\ U & \end{tikzcd} .$$

Proof (i) is immediate from the definition of $c(U)$.

(ii) Suppose A is a sheaf. If $B \to U$ is any object of $\mathcal{E}/U$, then the projection $B \times U \to B$ is an isomorphism, and so by (i) any subobject of B (in particular, $0 \rightarrowtail B$) is $c(U)$-dense. Thus the unique morphism $0 \to A$ extends to a unique morphism $B \to A$; hence there is also a unique morphism $B \to A \times U$. But this says that $A \times U$ is a terminal object of $\mathcal{E}/U$, so it is isomorphic to U. Conversely,

if $A \times U \cong U$ and we are given

$$\begin{tikzcd} B' \arrow[r, "f"] \arrow[d, tail, "m"] & A \\ B & \end{tikzcd}$$

where m is $c(U)$-dense, then the unique morphism $B \times U \to U \cong A \times U$ agrees with f on the intersection $B' \times U$ of B' and $B \times U$, and so can be combined with f to yield a unique extension of the latter to a morphism $B \to A$.

(iii) Since the functor $(-) \times U$ preserves pushouts (having a right adjoint), it is clear from (ii) that the pushout of the given diagram is a sheaf. On the other hand, for any sheaf B there is a unique morphism $U \to B$ (since $0 \rightarrowtail U$ is $c(U)$-dense), and so any morphism $A \to B$ factors uniquely through the pushout. □

Corollary 4.5.4 *For a subterminal object U of $\mathcal{E}$, the inclusion functor $\mathbf{sh}_{c(U)}(\mathcal{E}) \to \mathcal{E}$ creates all connected colimits which exist in $\mathcal{E}$.*

Proof This is immediate from 4.5.3(ii), since the functor $(-) \times U$ preserves colimits and the colimit of a connected diagram whose vertices are all U is (isomorphic to) U. □

Example 4.5.5 Reverting to the case of a functor category $[\mathcal{C}, \mathbf{Set}]$, we noted in 4.5.2 that subterminal objects of such a category correspond to cosieves in $\mathcal{C}$. If $\mathcal{R}$ is such a cosieve, then it is easy to see that a functor $F: \mathcal{C} \to \mathbf{Set}$ satisfies the condition that $F(a) \cong 1$ for all $a \in \mathrm{ob}\,\mathcal{R}$ iff it is isomorphic to the right Kan extension of its restriction to the full subcategory $\mathcal{S}$ on all objects *not* in $\mathcal{R}$. So by 4.5.3(ii) we may identify the closed subtopos of $[\mathcal{C}, \mathbf{Set}]$ complementary to $[\mathcal{R}, \mathbf{Set}]$ with $[\mathcal{S}, \mathbf{Set}]$. Note also that the condition that $\mathcal{R}$ should be a cosieve is equivalent to saying that $\mathcal{S}$ is a sieve, i.e. satisfies the dual condition that it contains any morphism whose codomain lies in $\mathcal{S}$. Thus we have a natural (order-preserving) bijection between closed subtoposes of $[\mathcal{C}, \mathbf{Set}]$ and open subtoposes of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, for any $\mathcal{C}$. (However, this identification depends on the Booleanness of $\mathbf{Set}$, as is clear from the use of complementation which we made in proving it: the analogous result for internal categories in a general topos is false.)

There is a close relationship between the notions of open and closed subtopos and the glueing construction of 2.1.12. If $F: \mathcal{E} \to \mathcal{F}$ is a cartesian functor between toposes, we observed in 4.1.12 that we have geometric morphisms

$$\mathcal{E} \xrightarrow{p} \mathbf{Gl}(F) \xleftarrow{q} \mathcal{F}$$

whose inverse images are the two projections; and from the descriptions of the direct images which we gave there, it is clear that they are both inclusions.

Further, p is an open inclusion, since we observed after 2.1.12 that its inverse image functor is logical; the corresponding subterminal object of $\mathbf{Gl}(F)$ is easily seen to be $(1, 0, 0 \to 1)$. And if we form the closed local operator corresponding to this subterminal object, it follows at once from 4.5.3(ii) that the sheaves for it are exactly those (A, B, f) for which $A \cong 1$, i.e. (up to isomorphism) exactly the objects in the image of q_*. So our original toposes $\mathcal{E}$ and $\mathcal{F}$ sit inside $\mathbf{Gl}(F)$ as complementary open and closed subtoposes; moreover, as we saw in 4.1.12, the original cartesian functor F may be recovered as the composite q^*p_*. (The composite p^*q_*, on the other hand, is simply the constant functor with value 1.) In the converse direction, we have

Proposition 4.5.6 *Let $\mathcal{E}$ be a topos, and U a subterminal object of $\mathcal{E}$. Let $\mathcal{L}$ and $\mathcal{M}$ be the open and closed subtoposes corresponding to U, and write $p \colon \mathcal{L} \to \mathcal{E}$, $q \colon \mathcal{M} \to \mathcal{E}$ for the inclusions. Then $\mathcal{E}$ is equivalent to the topos obtained by glueing along the composite $q^*p_* \colon \mathcal{L} \to \mathcal{M}$.*

Proof It is clear that p and q may be combined into a single geometric morphism $f \colon \mathcal{L} \times \mathcal{M} \to \mathcal{E}$, whose inverse image sends an object A to (p^*A, q^*A), and whose direct image sends (B, C) to $p_*B \times q_*C$. Moreover, from 4.5.3(i) we know that the monomorphisms in $\mathcal{E}$ which are dense for $o(U)$ are precisely those which are closed for $c(U)$, and in particular that any monomorphism which is dense for both local operators is an isomorphism. Hence $A \mapsto (p^*A, q^*A)$ preserves properness of subobjects, and so by 1.2.4 it is conservative, i.e. f is a surjection. So it suffices to show that the comonad on $\mathcal{L} \times \mathcal{M}$ induced by $(f^* \dashv f_*)$ is isomorphic to that induced by q^*p_*, as described in 4.2.4(c). But, for any object (B, C) of $\mathcal{L} \times \mathcal{M}$, we have $p^*p_*B \cong B$ and $q^*q_*C \cong C$, since p and q are inclusions; and since C is a sheaf for the closed local operator, we have $C^U \cong (C \times U)^U \cong U^U \cong 1$ by 4.5.3(ii), i.e. $p^*q_*C \cong 1$. So we see that

$$
\begin{aligned}
f^*f_*(B, C) &= (p^*(p_*B \times q_*C), q^*(p_*B \times q_*C)) \\
&\cong (p^*p_*B \times p^*q_*C, q^*p_*B \times q^*q_*C) \\
&\cong (B, q^*p_*B \times C);
\end{aligned}
$$

it is now easy to verify that this isomorphism of functors is an isomorphism of comonads, i.e. that it commutes with the counits and comultiplications. □

It will be observed that we have already encountered a particular instance of (the quasitopos version of) Proposition 4.5.6, in Proposition 2.6.7.

We have seen in 4.5.1 that the associated sheaf functor for a local operator is cartesian closed iff it is logical. However, it can preserve Ω without being logical; we next investigate the local operators for which this happens. (We shall discuss geometric morphisms whose *direct* images preserve Ω in the next section.)

First we need

Lemma 4.5.7 *For any local operator j on a topos $\mathcal{E}$, Ω_j is an 'internal exponential ideal' in Ω, i.e. the composite*

$$\Omega \times \Omega_j \rightarrowtail \Omega \times \Omega \xrightarrow{\Rightarrow} \Omega$$

factors through $\Omega_j \rightarrowtail \Omega$.

Proof By definition, the composite classifies the equalizer $E \rightarrowtail \Omega \times \Omega_j$ of

$$\Omega \times \Omega_j \rightarrowtail \Omega \times \Omega \underset{\wedge}{\overset{\pi_1}{\rightrightarrows}} \Omega;$$

so we have to show that E is a closed subobject. But given a morphism $(\phi, \psi): A \to \Omega \times \Omega_j$, corresponding to a pair of subobjects $A' \rightarrowtail A$, $A'' \rightarrowtail A$ of which the second is closed, we see that (ϕ, ψ) factors through E iff $A' \leq A''$ in $\mathrm{Sub}(A)$, iff $c_j(A') \leq A''$ in the lattice of closed subobjects of A. Since binary intersections of closed subobjects are closed by 4.3.3(iii), we deduce that $E \rightarrowtail \Omega \times \Omega_j$ is also the equalizer of

$$\Omega \times \Omega_j \xrightarrow{\bar{j} \times 1} \Omega_j \times \Omega_j \underset{\wedge}{\overset{\pi_1}{\rightrightarrows}} \Omega_j$$

where $\bar{j}: \Omega \twoheadrightarrow \Omega_j$ is the epic part of the image factorization of j; but Ω_j is a j-sheaf by 4.3.7, so this equalizer is closed by 4.3.6(a)(iv). □

Proposition 4.5.8 *Let j be a local operator on a topos $\mathcal{E}$. The following conditions are equivalent:*

(i) *The associated sheaf functor $L: \mathcal{E} \to \mathbf{sh}_j(\mathcal{E})$ preserves the subobject classifier.*

(ii) *The canonical monomorphism $\Omega_j \rightarrowtail \Omega$ is j-dense.*

(iii) *For any $\phi: A \to \Omega$, the equalizer of ϕ and $j\phi$ is a j-dense subobject of A.*

(iv) *Every monomorphism in $\mathcal{E}$ may be factored (not necessarily uniquely) as a j-closed monomorphism followed by a j-dense one.*

(v) *j commutes with implication, i.e. the diagram*

$$\begin{tikzcd} \Omega \times \Omega \arrow[r, "\Rightarrow"] \arrow[d, "j \times j"] & \Omega \arrow[d, "j"] \\ \Omega \times \Omega \arrow[r, "\Rightarrow"] & \Omega \end{tikzcd}$$

commutes.

(vi) *The diagram*

$$\begin{tikzcd} \Omega \arrow[r, "{(j,1)}"] \arrow[d] & \Omega \times \Omega \arrow[r, "\Rightarrow"] & \Omega \arrow[d, "j"] \\ 1 \arrow[rr, "\top"] & & \Omega \end{tikzcd}$$

commutes.

Proof (i) $\Leftrightarrow$ (ii): Let $\tau: L(\Omega) \to \Omega_j$ be the classifying map of $L(\top)$; then we have a diagram

$$\begin{tikzcd} J \arrow[r] \arrow[d, rightarrowtail] & L(1) \arrow[r] \arrow[d, rightarrowtail, "L(\top)"] & 1 \arrow[d, rightarrowtail] \\ \Omega \arrow[r, "\eta_\Omega"] & L(\Omega) \arrow[r, "\tau"] & \Omega_j \end{tikzcd}$$

in which the right-hand square is a pullback by definition, and the left-hand square is a pullback by the remarks preceding 4.3.2. It follows that the composite $\tau\eta_\Omega$ coincides with $\bar{j}$; hence if we identify Ω_j with $L(\Omega_j)$ τ becomes identified with $L(\bar{j})$. In particular, τ is an isomorphism iff $L(\bar{j})$ is an isomorphism; but since $\bar{j}$ is split epic, this holds iff $L(\Omega_j \rightarrowtail \Omega)$ is an isomorphism, i.e. iff $\Omega_j \rightarrowtail \Omega$ is j-dense.

(ii) $\Rightarrow$ (iii): The equalizer of $j\phi$ and ϕ is the pullback of $\Omega_j \rightarrowtail \Omega$ along ϕ; so it is dense if $\Omega_j \rightarrowtail \Omega$ is.

(iii) $\Rightarrow$ (iv): If ϕ is the classifying map of $A' \rightarrowtail A$, then the equalizer $A'' \rightarrowtail A$ of ϕ and $j\phi$ is the largest subobject which contains A' as a j-closed subobject. So this is immediate.

(iv) $\Rightarrow$ (ii): Applying (iv) to the generic subobject $\top: 1 \rightarrowtail \Omega$, we obtain a j-dense subobject $B \rightarrowtail \Omega$ containing $\top$ as a j-closed subobject. But then $B \rightarrowtail \Omega$ must factor through $\Omega_j \rightarrowtail \Omega$, so the latter is also dense.

(iii) $\Rightarrow$ (v): Since both ways round the given diagram factor through $\Omega_j \rightarrowtail \Omega$ (the lower one by an application of Lemma 4.5.7), it suffices (since Ω_j is j-separated) to find a j-dense subobject $E \rightarrowtail \Omega \times \Omega$ equalizing them. But (iii) implies that we may take E to be the intersection of the equalizers of the pairs $(j\pi_1, \pi_1)$, $(j\pi_2, \pi_2)$ and $(j\Rightarrow, \Rightarrow)$.

(v) $\Rightarrow$ (vi) follows from composing the two ways round the diagram of (v) with $(j,1): \Omega \to \Omega \times \Omega$, since the composite $\Rightarrow\Delta: \Omega \to \Omega \times \Omega \to \Omega$ factors through $\top$.

(vi) $\Rightarrow$ (iii): Since any subobject is contained in its closure, we have $\wedge(j\phi, \phi) = \phi$ for any $\phi: A \to \Omega$, and hence the composite

$$A \xrightarrow{\phi} \Omega \xrightarrow{(j,1)} \Omega \times \Omega \xrightarrow{\Rightarrow} \Omega$$

classifies the equalizer of $j\phi$ and ϕ. Condition (vi) says immediately that this equalizer is j-dense. □

Example 4.5.9 Let $\neg\colon \Omega \to \Omega$ be the Heyting negation map, i.e. the classifying map of $\bot\colon 1 \rightarrowtail \Omega$. It is straightforward to verify that the composite $\neg\neg$ is a local operator, i.e. that it satisfies the conditions of 4.4.1. Moreover, it satisfies the conditions of 4.5.8: to see this, observe that for any element x of a Heyting algebra H, we have $x \leq (\neg\neg x \Rightarrow x)$ and $\neg x \leq (\neg\neg x \Rightarrow x)$ (the latter since $(\neg x \wedge \neg\neg x) = \bot \leq x$), and so $(\neg\neg x \Rightarrow x) \geq (x \vee \neg x)$; hence $\neg\neg(\neg\neg x \Rightarrow x) \geq \neg\neg(x \vee \neg x) = \top$. But this is just the statement that the diagram in (vi) of 4.5.8 commutes. Alternatively, we could use condition (iv): given a subobject $A' \rightarrowtail A$, if we set $A'' = A' \cup \neg A'$, then $A' \rightarrowtail A''$ is $\neg\neg$-closed (since it is complemented) and $A'' \rightarrowtail A$ is $\neg\neg$-dense (cf. the proof of 1.4.14).

We note that the subtopos $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ is Boolean; for if A is any $\neg\neg$-sheaf, its subobjects in $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ are its $\neg\neg$-closed subobjects in $\mathcal{E}$, and these form a Boolean algebra. It is easy to see that it is not an open subtopos in general; for example, if X is a T_0-space (such as $\mathbb{R}$) in which no nonempty open subspace is discrete, then $\mathbf{sh}_{\neg\neg}(\mathbf{Sh}(X))$ cannot be open. We shall have more to say about Boolean subtoposes in 4.5.21 below.

We write $\mathbf{Lop}(\mathcal{E})$ for the class of all local operators on a topos $\mathcal{E}$ (note that it is a set if $\mathcal{E}$ is locally small). $\mathbf{Lop}(\mathcal{E})$ carries a natural partial order, defined by $j_1 \leq j_2$ iff $\wedge(j_1, j_2) = j_1$; this is equivalent to saying that $J_1 \leq J_2$ in $\mathrm{Sub}(\Omega)$, or that $\Omega_{j_2} \leq \Omega_{j_1}$, or that $\mathbf{sh}_{j_2}(\mathcal{E}) \subseteq \mathbf{sh}_{j_1}(\mathcal{E})$ as subcategories of $\mathcal{E}$ (the more dense monomorphisms we have, the more conditions an object has to satisfy to be a sheaf). We shall see eventually that $\mathbf{Lop}(\mathcal{E})$ is a Heyting algebra; for the moment, we note

Lemma 4.5.10 *The partial ordering* $\mathbf{Lop}(\mathcal{E})$ *has greatest and least elements, and binary meets.*

Proof The composite $\Omega \to 1 \xrightarrow{\top} \Omega$ is a local operator (it is the open local operator associated with the subterminal object 0), and is clearly the greatest element of $\mathbf{Lop}(\mathcal{E})$. Similarly, the least element is the local operator 1_Ω. The binary meets are constructed 'pointwise' – that is, if j and k are local operators, so is the composite

$$\Omega \xrightarrow{(j,k)} \Omega \times \Omega \xrightarrow{\wedge} \Omega\,.$$

It is straightforward to verify that this composite satisfies the commutative diagrams of 4.4.1; alternatively, one may verify that if c and d are universal closure operations on a cartesian category, so is the mapping $A' \mapsto c(A') \cap d(A')$. □

We further note that if U is a subterminal object, then $(o(U) \wedge c(U)) = 1_\Omega$, as claimed earlier; for in any Heyting algebra we have

$$\begin{aligned}(u \Rightarrow x) \wedge (u \vee x) &= ((u \Rightarrow x) \wedge u) \vee ((u \Rightarrow x) \wedge x) \\ &= (u \wedge x) \vee x = x\end{aligned}$$

by the identities of 1.5.11 and the distributive law.

A problem which frequently arises is that of constructing the smallest local operator (in the ordering defined above) for which all the monomorphisms in a given class are dense. We next present a particularly elegant solution to this problem, due to A. Joyal. First we need a lemma characterizing the class of dense monomorphisms for a local operator. We introduce the notation $\Xi(D)$ for the class of monomorphisms whose classifying maps factor through a given subobject D of Ω; thus if j is a local operator, $\Xi(J)$ is the class of j-dense monomorphisms, and $\Xi(\Omega_j)$ is the class of j-closed monomorphisms.

Lemma 4.5.11 *Let $D \rightarrowtail \Omega$ be a subobject of Ω in a topos $\mathcal{E}$. Then*

(i) *$\Xi(D)$ is stable under pullbacks.*

(ii) *If $\top: 1 \to \Omega$ factors through D (equivalently, if $\Xi(D)$ contains all isomorphisms), then $\Xi(D)$ is stable under pushouts.*

(iii) *The classifying map of $D \rightarrowtail \Omega$ is a local operator iff $\Xi(D)$ contains all isomorphisms and satisfies the condition ($mn \in \Xi(D) \Leftrightarrow m \in \Xi(D)$ and $n \in \Xi(D)$) for all composable pairs of monomorphisms m, n.*

Proof (i) is immediate since if m is classified by ϕ then $f^*(m)$ is classified by ϕf.

(ii) Let

$$\begin{tikzcd} A' \arrow[r, "f'"] \arrow[d, tail, "m"] & B' \arrow[d, tail, "n"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

be a pushout diagram with $m \in \Xi(D)$. By 2.4.3, the square is also a pullback; so if $\phi: B \to \Omega$ is the classifying map of n, then ϕf classifies m, and so factors through D. But ϕn classifies $n^*(n) \cong 1_{B'}$, so it too factors through D; the result follows from the universal property of pushouts.

(iii) If the classifying map $d: \Omega \to \Omega$ of $D \rightarrowtail \Omega$ is a local operator, then $\Xi(D)$ is the class of d-dense monomorphisms, and it satisfies the given conditions by 4.3.3. Conversely, suppose the conditions are satisfied. It suffices by 4.4.2 to show that the unary operation (c, say) on subobjects which sends the subobject classified by ϕ to that classified by $d\phi$ is a universal closure operation. But since $\Xi(D)$ contains isomorphisms, it is clear that we have $c(A') \geq A'$; and in fact

$c(A')$ is the largest A'' in $\mathrm{Sub}(A)$ for which $A' \rightarrowtail A''$ is in $\Xi(D)$. Now the second condition on $\Xi(D)$ implies that the composite $A' \rightarrowtail c(A') \rightarrowtail c(c(A'))$ is in $\Xi(D)$, so $c(c(A')) \leq c(A')$, i.e. c is idempotent. To show that c is order-preserving, consider two subobjects $A' \rightarrowtail A$ and $A'' \rightarrowtail A$ with $A' \leq A''$. Form the diagram

$$\begin{tikzcd} c(A') \cap c(A'') \arrow[r, tail] \arrow[d, tail, "m"] & c(A'') \arrow[d, tail, "n"] \\ c(A') \arrow[r, tail] & c(A') \cup c(A'') \end{tikzcd} .$$

Now since $A' \leq A'' \leq c(A'')$, the monomorphism $A' \rightarrowtail c(A')$ factors through m, so $m \in \Xi(D)$. And the square is a pushout by 1.4.3, so by (ii) we have $n \in \Xi(D)$. But now the composite $A'' \rightarrowtail c(A'') \rightarrowtail c(A') \cup c(A'')$ is in $\Xi(D)$, so n must be an isomorphism; that is, $c(A') \leq c(A'')$. Finally, it is clear from the form of the definition that c is stable under pullback. □

Next we define a binary relation θ on $\mathrm{Sub}(A)$, for each object A of $\mathcal{E}$, by saying that $\theta(A', A'')$ holds iff $(A' \Rightarrow A'') \cong A''$. We note that this relation always holds if A' is dense for some local operator j and A'' is closed for the same local operator; for we have $A' \cap (A' \Rightarrow A'') \leq A'' \leq (A' \Rightarrow A'')$, and $A' \cap (A' \Rightarrow A'')$ is j-dense in $(A' \Rightarrow A'')$. Note also that $\theta(A', A')$ holds iff A' is the whole of A, since $(A' \Rightarrow A') \cong A$. It is clear that, if we define $\Theta \rightarrowtail \Omega \times \Omega$ to be the equalizer of $\Rightarrow$ and π_2, then a morphism $(\phi, \psi)\colon A \to \Omega \times \Omega$ factors through Θ iff the subobjects classified by ϕ and ψ satisfy the relation θ. Hence in particular if j is a local operator, then $J \times \Omega_j \leq \Theta$ in $\mathrm{Sub}(\Omega \times \Omega)$.

We now use Θ to define an 'internal Galois connection' on $\mathrm{Sub}_{\mathcal{E}}(\Omega)$, as follows: if $D \rightarrowtail \Omega$, we define $D^r \rightarrowtail \Omega$ to be

$$\forall_{\pi_2}((\pi_1^*(D) \Rightarrow \Theta) \rightarrowtail \Omega \times \Omega)$$

where π_1 and π_2, as usual, denote the product projections $\Omega \times \Omega \to \Omega$. (Equivalently, D^r is the unique largest subobject of Ω such that $D \times D^r \leq \Theta$ in $\mathrm{Sub}(\Omega \times \Omega)$.) And we define $D^l \rightarrowtail \Omega$ by the same formula with π_1 and π_2 interchanged. It is straightforward to verify that a monomorphism $m\colon A' \rightarrowtail A$ belongs to $\Xi(D^r)$ iff, for every $f\colon B \to A$ and every $n\colon B' \rightarrowtail B$ in $\Xi(D)$, the relation $\theta(n, f^*(m))$ holds.

Proposition 4.5.12 (i) *The mappings $D \mapsto D^r$ and $D \mapsto D^l$ form a Galois connection from* $\mathrm{Sub}(\Omega)$ *to itself; that is, they are order-reversing and adjoint to each other on the right.*

(ii) *If j is a local operator on $\mathcal{E}$, then $J^r \cong \Omega_j$ and $(\Omega_j)^l \cong J$.*

(iii) *A subobject $D \rightarrowtail \Omega$ satisfies $D^{rl} \cong D$ iff its classifying map is a local operator on $\mathcal{E}$.*

Proof (i) is immediate from the description of D^r given above, and the corresponding description of D^l.

(ii) We have already observed that if j is a local operator then $J \times \Omega_j \leq \Theta$, so we have $\Omega_j \leq J^r$ and $J \leq (\Omega_j)^l$.

To show that we have equality in the first case, suppose $(A' \rightarrowtail A) \in \Xi(J^r)$. Let $A'' \rightarrowtail A$ be the j-closure of A'; then $A' \rightarrowtail A''$ is the pullback of $A' \rightarrowtail A$ along $A' \rightarrowtail A''$; so it too belongs to $\Xi(J^r)$. But it also belongs to $\Xi(J)$, since it is j-dense; so it stands in the relation θ to itself, and hence must be an isomorphism, i.e. $A' \rightarrowtail A$ is closed. Applying this argument to the subobject of J^r classified by the monomorphism $J^r \rightarrowtail \Omega$, we deduce that the latter must factor through Ω_j.

Similarly in the second case, suppose that $(B' \rightarrowtail B) \in \Xi((\Omega_j)^l)$. Let $B'' \rightarrowtail B$ be the j-closure of B'; then since it belongs to $\Xi(\Omega_j)$ we have $\theta(B', B'')$; that is, $B'' \cong (B' \Rightarrow B'') \cong B$ since $B' \leq B''$. So $B' \rightarrowtail B$ is j-dense. Once again, applying this in the generic case $B = (\Omega_j)^l$ yields $(\Omega_j)^l \leq J$.

(iii) If the classifying map of $D \rightarrowtail \Omega$ is a local operator, then $D^{rl} \cong D$ follows from (ii). Conversely, suppose $D^{rl} \cong D$; we shall show that the class $\Xi(D)$ satisfies the conditions of 4.5.11(iii).

If $m: A' \rightarrowtail A$ is an isomorphism, then it is clear that $\theta(A', A'')$ holds for any A''; hence $m \in \Xi(E^l)$ for any E, and in particular $m \in \Xi(D^{rl}) = \Xi(D)$. Now suppose we have a composable pair $n: A'' \rightarrowtail A'$, $m: A' \rightarrowtail A$ of monomorphisms in $\Xi(D)$. Then for any $p: A''' \rightarrowtail A$ in $\Xi(D^r)$, we have $\theta(m, p)$ and $\theta(n, m^*(p))$, from which it is easy to deduce that $\theta(mn, p)$ holds. And this remains true after pullback along any $f: B \to A$, so we deduce that $mn \in \Xi(D^{rl}) = \Xi(D)$. Conversely if $mn \in \Xi(D)$, then it is clear that $n \in \Xi(D)$, since $n \cong m^*(mn)$; and for any $p: A''' \rightarrowtail A$ in $\Xi(D^r)$, we can deduce $\theta(m, p)$ from $\theta(mn, p)$. Once again, this remains true after pullback, so we have $m \in \Xi(D^{rl}) = \Xi(D)$. □

Corollary 4.5.13 *Let $D \rightarrowtail \Omega$ be a subobject of Ω in a topos. Then*

(i) *There is a unique smallest local operator j such that all the morphisms in $\Xi(D)$ are j-dense.*

(ii) *There is a unique largest local operator k such that all the morphisms in $\Xi(D)$ are k-closed.*

Proof (i) We take j to be the classifying map of $D^{rl} \rightarrowtail \Omega$; the result is immediate from 4.5.12(i) and (iii).

(ii) Similarly, we take k to be the classifying map of $D^l \rightarrowtail \Omega$. □

Examples 4.5.14 (*a*) We may now verify that $\mathbf{Lop}(\mathcal{E})$ has binary joins, as well as the binary meets constructed earlier; we simply define $j_1 \vee j_2$ to be the smallest local operator for which all monomorphisms in $\Xi(J_1 \cup J_2)$ are dense. (Note in passing that the closed monomorphisms for $j_1 \vee j_2$ are exactly those which are closed for both j_1 and j_2, since $\Omega_{j_1 \vee j_2} \cong (J_1 \cup J_2)^r \cong (J_1^r \cap J_2^r)$. It follows that if j_1 and j_2 are the open and closed local operators associated with

a subterminal object U, then $j_1 \vee j_2$ is the top element of $\mathbf{Lop}(\mathcal{E})$; for we saw in 4.5.3(i) that every j_2-closed monomorphism is j_1-dense, and so a monomorphism which is closed for both local operators must be an isomorphism.)

(*b*) Let $m: A' \rightarrowtail A$ be a single monomorphism in a topos $\mathcal{E}$. If we apply 4.5.13(i) to the image $I \rightarrowtail \Omega$ of the classifying map of m, we obtain the smallest local operator on $\mathcal{E}$ for which m is dense. (Dually, we may obtain the largest local operator for which m is closed, using 4.5.13(ii).) For example, if U is a subterminal object of $\mathcal{E}$, it may easily be verified that $o(U)$ is the smallest local operator for which $U \rightarrowtail 1$ is dense, and $c(U)$ is the smallest for which $0 \rightarrowtail U$ is dense.

(*c*) Let $f: A \to B$ be a morphism in a topos. By applying (*b*) to the image of f (resp. the diagonal map $A \rightarrowtail A \times_B A$), we may obtain the smallest local operator for which $L(f)$ is epic (resp. monic), where L is the associated sheaf functor, as usual. And by taking the join of these two, we may obtain the smallest local operator for which $L(f)$ is an isomorphism.

(*d*) Similarly, by applying the dual of (*b*) to the diagonal map $A \rightarrowtail A \times A$ of an object A of $\mathcal{E}$, we may obtain the largest local operator on $\mathcal{E}$ for which A is separated, by 4.3.6(iv).

(*e*) Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism, j a local operator on $\mathcal{E}$. Then there is a unique smallest local operator k on $\mathcal{F}$ such that the composite $\mathbf{sh}_k(\mathcal{F}) \to \mathcal{F} \to \mathcal{E}$ factors through $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$. For such a factorization exists, by 4.3.12, iff $f^*(m)$ is k-dense for every j-dense monomorphism m; so we simply have to apply 4.5.13(i) to the image of the composite

$$f^*(J) \rightarrowtail f^*(\Omega_{\mathcal{E}}) \xrightarrow{\tau} \Omega_{\mathcal{F}}$$

where τ is the classifying map of $f^*(\top)$. The resulting local operator k is sometimes called the *pullback local operator*, since it is easy to see using 4.3.11 that the square

$$\begin{tikzcd} \mathbf{sh}_k(\mathcal{F}) \arrow[r] \arrow[d] & \mathbf{sh}_j(\mathcal{E}) \arrow[d] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

is a pullback in $\mathfrak{Top}$. For, given a topos $\mathcal{G}$, a cone over the right and bottom edges of this square with vertex $\mathcal{G}$ corresponds to a geometric morphism $g: \mathcal{G} \to \mathcal{F}$ such that $g^* f^*$ maps j-dense monomorphisms to isomorphisms; but if it does so, then the local operator k' on $\mathcal{F}$ which corresponds to the image of g (in the sense of 4.2.10) must satisfy $k' \geq k$, and so g factors through $\mathbf{sh}_k(\mathcal{F}) \to \mathcal{F}$.

(*f*) We have already seen that $\mathbf{Lop}(\mathcal{E})$ is a lattice; we now show that it is a Heyting algebra (and hence in particular a distributive lattice). To do this, let j_1 and j_2 be local operators, and define j_3 to be the largest local operator such that every monomorphism which is both j_1-dense and j_2-closed is j_3-closed, i.e.

set $J_3 = (J_1 \cap \Omega_{j_2})^l$. We shall show that $j_3 = (j_1 \Rightarrow j_2)$ in $\mathbf{Lop}(\mathcal{E})$. Clearly, if k is any local operator such that $k \wedge j_1 \leq j_2$, then any $A' \rightarrowtail A$ which is both j_1-dense and j_2-closed must be k-closed, for the inclusion $A' \rightarrowtail A''$ of A' in its k-closure would be both k-dense and j_1-dense (and hence j_2-dense), and also j_2-closed. Hence any such k satisfies $k \leq j_3$; we thus have to prove that j_3 itself satisfies $j_3 \wedge j_1 \leq j_2$. Suppose $m: A' \rightarrowtail A$ is both j_3-dense and j_1-dense. If $n: A'' \rightarrowtail A$ is any j_2-closed monomorphism, let $p: A''' \rightarrowtail A$ be its j_1-closure, and $q: A'' \rightarrowtail A'''$ the factorization of n through p. Then we have $\theta(m, p)$ since m is j_1-dense and p is j_1-closed, and $\theta(p^*(m), q)$ since $p^*(m)$ is j_3-dense and $q = p^*(n)$ is both j_1-dense and j_2-closed. From these it follows that we have $\theta(m, pq)$; and the same argument works if we first pull back m along some $f: B \to A$, so $m \in \Xi((\Omega_{j_2})^l) = \Xi(J_2)$.

We saw in 4.5.14(d) that there is a largest local operator for which a given object A is separated. The same result holds true if we replace 'separated' by 'a sheaf', but it does not seem possible to derive it directly from 4.5.13 (unless A happens to be subterminal, in which case we may take the largest local operator for which $A \rightarrowtail 1$ is closed – cf. 4.5.21 below).

Proposition 4.5.15 *For any object A of a topos $\mathcal{E}$, there is a largest local operator on $\mathcal{E}$ for which A is a sheaf.*

Proof Let $\psi: A \times \Omega \to \tilde{A}$ represent the partial map

$$\begin{tikzcd} A \arrow[r, "1_A"] \arrow[d, tail, "1_A \times \top"] & A \\ A \times \Omega & \end{tikzcd}$$

and let $\tau: \tilde{A} \to \Omega$ classify $A \rightarrowtail \tilde{A}$. Note that ψ is the case $J = \Omega$ of the morphism considered in 4.4.10; so we want to find the largest subobject J of Ω such that the pullback of $\psi: \Omega^* A \to \tau$ along $J \rightarrowtail \Omega$ is an isomorphism. To do this, form the pullback

$$\begin{tikzcd} W \arrow[rr, "(v,w)"] \arrow[d, tail, "u"] & & A \times \Omega \arrow[d, tail, "\{\}"] \\ \tilde{A} \arrow[r, "\{\}"] & P\tilde{A} \arrow[r, "P\psi"] & P(A \times \Omega) \end{tikzcd} \quad .$$

Now a map $f\colon B \to \tilde{A}$ factors through u iff there exists $g\colon B \to A$ such that

$$\begin{tikzcd}
B \arrow[rr, "1_B"] \arrow[d, tail, "(1_B, g, \tau f)"] & & B \arrow[d, tail, "(1_B, f)"] \\
B \times A \times \Omega \arrow[rr, "1_B \times \psi"] & & B \times \tilde{A}
\end{tikzcd}$$

is a pullback, since the composite $P\psi \circ \{\} \circ f$ names the pullback of the graph of f along $1_B \times \psi$ (cf. 2.2.2(i)). But this holds iff, for every $h\colon C \to B$, the composite gh is the unique morphism $C \to A$ extending the partial map $C \rightharpoonup A$ represented by fh. Now if we define $J \rightarrowtail \Omega$ to be $\forall_\tau(W \rightarrowtail \tilde{A})$, then a morphism $\chi\colon B \to \Omega$ factors through J iff, for all $h\colon C \to B$, the subobject $C' \rightarrowtail C$ classified by χh has the property that each morphism $C' \to A$ extends uniquely to a morphism $C \to A$. It follows easily that the class $\Xi(J)$ satisfies the conditions of 4.5.11(iii), so that the classifying map of $J \rightarrowtail \Omega$ is a local operator; and it is clearly the largest local operator for which A is a sheaf. □

Corollary 4.5.16 *If j_1 and j_2 are two local operators on $\mathcal{E}$, then $\mathbf{sh}_{(j_1 \vee j_2)}(\mathcal{E}) = \mathbf{sh}_{j_1}(\mathcal{E}) \cap \mathbf{sh}_{j_2}(\mathcal{E})$ and $\mathbf{sep}_{j_1 \vee j_2}(\mathcal{E}) = \mathbf{sep}_{j_1}(\mathcal{E}) \cap \mathbf{sep}_{j_2}(\mathcal{E})$.*

Proof An object A of $\mathcal{E}$ is a sheaf for $j_1 \vee j_2$ iff the local operator j constructed in 4.5.15 satisfies $j \geq (j_1 \vee j_2)$, iff A is a sheaf for both j_1 and j_2. The second assertion follows similarly from 4.5.14(d). □

In general, there is no simple formula for the join of two local operators, comparable to that for their meet which we gave in 4.5.10. However, the following is sometimes useful.

Lemma 4.5.17 *Let j and m be two local operators on a topos $\mathcal{E}$. The following conditions are equivalent:*

(i) *The join of j and m in $\mathbf{Lop}(\mathcal{E})$ is the composite mj.*

(ii) *mj is a local operator.*

(iii) *mj is idempotent.*

(iv) *$mj \geq jm$ in the canonical partial ordering on $\mathcal{E}(\Omega, \Omega)$.*

Proof Clearly (i) implies (ii), but (ii) implies (i) since we have $1_\Omega \leq j$, $1_\Omega \leq m$ and hence $m \leq mj$, $j \leq mj$ in $\mathcal{E}(\Omega, \Omega)$, and

$$mj \leq (j \vee m)(j \vee m) = (j \vee m)\,.$$

(ii) $\Rightarrow$ (iii) is again immediate, and the converse holds since it easily seen that mj always satisfies the first and third commutative diagrams of 4.4.1. (iii) implies (iv) since we have $mj = mjmj \geq 1_\Omega jm1_\Omega$; and (iv) implies (iii) since we have $mjmj \leq mmjj = mj$, the reverse inequality being trivial. □

The conditions of 4.5.17 are easily seen to hold if either m is an open local operator, or j is closed.

Although calculating the join of two local operators j and m as a local operator on $\mathcal{E}$ is in general a hard problem, it turns out that calculating the local operator k on $\mathbf{sh}_m(\mathcal{E})$ which corresponds to the subtopos $\mathbf{sh}_{j\vee m}(\mathcal{E}) = \mathbf{sh}_j(\mathcal{E}) \cap \mathbf{sh}_m(\mathcal{E})$ is often easy. The latter is of course the pullback of j along the inclusion $f\colon \mathbf{sh}_m(\mathcal{E}) \to \mathcal{E}$, in the sense of 4.5.14(e), so we know that it is the smallest local operator for which the corresponding $K \rightarrowtail \Omega_m$ contains the image of $f^*J \rightarrowtail f^*\Omega \to \Omega_m$. But in many cases it turns out that the classifying map of this image is already a local operator.

Lemma 4.5.18 *Let j and m be local operators on a topos $\mathcal{E}$, and assume one of the following:*

(i) *The pair (j, m) satisfies the conditions of* 4.5.17*;*
(ii) *m is closed;*
(iii) *m satisfies the conditions of* 4.5.8.

Then the classifying map of the m-closure of the image of the composite

$$J \rightarrowtail \Omega \xrightarrow{\bar{m}} \Omega_m$$

is a local operator k in $\mathbf{sh}_m(\mathcal{E})$, and it is the pullback of j (in the sense of 4.5.14(*e*)*) along the inclusion $f\colon \mathbf{sh}_m(\mathcal{E}) \to \mathcal{E}$.*

Proof The second assertion is immediate from the first, since the closure of the image of $J \to \Omega_m$ is exactly the image of the composite $f^*J \rightarrowtail f^*\Omega \to \Omega_m$ in $\mathbf{sh}_m(\mathcal{E})$. Let us write $K \rightarrowtail \Omega_m$ for this subobject: then it is easy to see that a (necessarily m-closed) monomorphism $B' \rightarrowtail B$ in $\mathbf{sh}_m(\mathcal{E})$ belongs to $\Xi(K)$ iff it can be factored in $\mathcal{E}$ as $B' \rightarrowtail B'' \rightarrowtail B$ where $B' \rightarrowtail B''$ is j-dense and $B'' \rightarrowtail B$ is m-dense. (Specifically, we take $B'' \rightarrowtail B$ to be the pullback along $B \to K$ of the m-dense monomorphism $I \rightarrowtail K$, where $I \rightarrowtail \Omega_m$ is the image of $J \to \Omega_m$.) We have to show that $\Xi(K)$ satisfies the conditions of 4.5.11(iii); but they are all trivial except for stability under composition, for which we argue differently in the three cases, as follows.

(i) The description just given says that $\Xi(K)$ consists of those monomorphisms $B' \rightarrowtail B$ whose classifying maps $\phi\colon B \to \Omega$ in $\mathcal{E}$ satisfy $mj\phi = \top B$. If $B' \rightarrowtail B$ is a composite of two such monomorphisms, then its classifying map satisfies $mjmj\phi = \top B$; but the conditions of 4.5.17 say that this is equivalent to $mj\phi = \top B$.

(ii) If m is a closed local operator, then it follows from 4.5.3(i) that any subobject containing an m-closed one is m-closed. So the description of $\Xi(K)$ simplifies still further: it consists of those monomorphisms in $\mathbf{sh}_m(\mathcal{E})$ which are j-dense. As such, it is clearly closed under composition.

(iii) Here we argue slightly differently: in this case the inverse image functor $f^*: \mathcal{E} \to \mathbf{sh}_m(\mathcal{E})$ preserves Ω, by 4.5.8(i), and the isomorphism $f^*\Omega \cong \Omega_m$ identifies $K \rightarrowtail \Omega_m$ with $f^*J \rightarrowtail f^*\Omega$. Since f^* preserves pullbacks, it also identifies the classifying map k of $K \rightarrowtail \Omega_m$ with $f^*(j)$; but the latter clearly inherits the conditions of 4.4.1 from j, so it is a local operator. □

Next, we investigate how the open and closed local operators sit inside the lattice $\mathbf{Lop}(\mathcal{E})$.

Lemma 4.5.19 (i) *The mapping* c: $\mathrm{Sub}_{\mathcal{E}}(1) \to \mathbf{Lop}(\mathcal{E})$ *is a lattice homomorphism, and has a right adjoint* ext.

(ii) *The mapping* o: $\mathrm{Sub}_{\mathcal{E}}(1)^{\mathrm{op}} \to \mathbf{Lop}(\mathcal{E})$ *is a lattice homomorphism, and has a left adjoint* int.

Proof (i) The fact that c preserves finite meets follows easily from 4.5.10 and the distributive law in Ω; it preserves finite joins by the remark after 4.5.17. Given j, we define $\mathrm{ext}(j)$ (the *exterior* of j) to be the j-closure of $0 \rightarrowtail 1$ (i.e. the subobject classified by $j\bot$). It is clear that $\mathrm{ext}(c(U)) \cong U$; and for any j we have $c(\mathrm{ext}(j)) \leq j$, since for any subobject $A' \rightarrowtail A$ we have both $A' \leq c_j(A')$ and $c_j(0 \rightarrowtail A) \cong (A \times \mathrm{ext}(j)) \leq c_j(A')$, where c_j is the universal closure operator corresponding to j. So the adjunction is established.

(ii) The fact that o is a lattice homomorphism may be verified directly as in (i), or alternatively deduced from (i) and the fact (which we have already observed) that $c(U)$ and $o(U)$ are complementary elements of $\mathbf{Lop}(\mathcal{E})$. For the adjunction, we take the *interior* $\mathrm{int}(j)$ of j to be $\Pi_J(d)$, where d: $1 \rightarrowtail J$ is the generic j-dense subobject (cf. 4.5.1). Now we have $U \leq \mathrm{int}(j)$ iff $J^*(U) \leq d$ in $\mathrm{Sub}(J)$; but this is equivalent to saying that every j-dense subobject is $o(U)$-dense, i.e. that $j \leq o(U)$. □

The local operator $c(\mathrm{ext}(j))$ is of course called the *closure* of j. We note that the inclusion $\mathbf{sh}_j(\mathcal{E}) \to \mathbf{sh}_{c(\mathrm{ext}(j))}(\mathcal{E})$ corresponding to the inequality $c(\mathrm{ext}(j)) \leq j$ has the property that its direct image preserves the initial object (equivalently, the local operator on $\mathbf{sh}_{c(\mathrm{ext}(j))}(\mathcal{E})$ which corresponds to it has exterior 0), since the initial object of $\mathbf{sh}_j(\mathcal{E})$ is simply $\mathrm{ext}(j)$. Of course, we call a subtopos or local operator *dense* if it has this property; we may thus conclude

Corollary 4.5.20 *Any geometric inclusion $\mathcal{E}' \to \mathcal{E}$ has a unique factorization $\mathcal{E}' \to \mathcal{E}'' \to \mathcal{E}$, where $\mathcal{E}' \to \mathcal{E}''$ is dense and $\mathcal{E}'' \to \mathcal{E}$ is closed.*

Proof The existence has been discussed above. For the uniqueness, let j and j' be the local operators on $\mathcal{E}$ corresponding to $\mathcal{E}'$ and $\mathcal{E}''$; then the denseness of the first inclusion is equivalent to the equality $\mathrm{ext}(j) = \mathrm{ext}(j')$, and so the closedness of the second implies that $j' = c(\mathrm{ext}(j))$. □

We note that the local operator $\neg\neg$ of 4.5.9 is dense. Indeed, it is the unique largest dense local operator on $\mathcal{E}$ (i.e. $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ is the smallest dense subtopos): for if j is a dense local operator, then c_j preserves disjointness of subobjects by

4.3.3(iii), and so $c_j(A') \cap \neg A' \leq c_j(A') \cap c_j(\neg A') = 0$, whence $c_j(A') \leq \neg\neg A'$ for any $A' \rightarrowtail A$.

We conclude this section by discussing the relation between local operators and Booleanness. For this purpose, we introduce a third family of local operators indexed by the subterminal objects of $\mathcal{E}$: given a subterminal object U, we define the *quasi-closed* local operator $q(U)$ associated with U to be the composite

$$\Omega \cong \Omega \times 1 \xrightarrow{1 \times (u,u)} \Omega \times \Omega \times \Omega \xrightarrow{\Rightarrow \times 1} \Omega \times \Omega \xrightarrow{\Rightarrow} \Omega$$

where u, as before, is the classifying map of $U \rightarrowtail 1$. (In particular, we note that $q(0)$ is the local operator $\neg\neg$ considered in 4.5.9.) It is again a straightforward exercise in the algebraic theory of Heyting algebras to verify that this map satisfies the commutative diagrams of 4.4.1. Further, we have $\mathrm{ext}(q(U)) = U = \mathrm{ext}(c(U))$, so we have a composite inclusion

$$\mathbf{sh}_{q(U)}(\mathcal{E}) \longrightarrow \mathbf{sh}_{c(U)}(\mathcal{E}) \longrightarrow \mathcal{E}$$

as in 4.5.20; it is easily seen that this factorization identifies $\mathbf{sh}_{q(U)}(\mathcal{E})$ with $\mathbf{sh}_{\neg\neg}(\mathbf{sh}_{c(U)}(\mathcal{E}))$, and hence in particular $\mathbf{sh}_{q(U)}(\mathcal{E})$ is Boolean.

Lemma 4.5.21 *Let j be a local operator on a topos $\mathcal{E}$. Then $\mathbf{sh}_j(\mathcal{E})$ is Boolean iff $j = q(U)$ for some subterminal object U of $\mathcal{E}$.*

Proof One direction follows from the remarks above. Conversely, suppose $\mathbf{sh}_j(\mathcal{E})$ is Boolean. By factoring the inclusion $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$ as in 4.5.20, we may reduce to the case when j is dense; in this case, we have to prove that $j = \neg\neg$. Now it is easy to see that a composite of two inclusions is dense iff each of the two factors is dense; so, for any dense local operator j, the smallest dense subtopos $\mathbf{sh}_{\neg\neg}(\mathbf{sh}_j(\mathcal{E}))$ of $\mathbf{sh}_j(\mathcal{E})$ coincides with the smallest dense subtopos $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ of $\mathcal{E}$. But if $\mathbf{sh}_j(\mathcal{E})$ is Boolean, its double-negation operator is the identity; so it coincides with $\mathbf{sh}_{\neg\neg}(\mathbf{sh}_j(\mathcal{E}))$ and hence with $\mathbf{sh}_{\neg\neg}(\mathcal{E})$. Hence $j = \neg\neg$. □

We note that, since $q(U)$ is the largest local operator for which the inclusion $\mathbf{sh}_{q(U)}(\mathcal{E}) \to \mathbf{sh}_{c(U)}(\mathcal{E})$ is dense, it may also be characterized as the largest local operator on $\mathcal{E}$ for which $U \rightarrowtail 1$ is closed (equivalently, for which U is a sheaf).

Proposition 4.5.22 *For a topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is Boolean.*

(ii) *Every subtopos of $\mathcal{E}$ is Boolean.*

(iii) *Every subtopos of $\mathcal{E}$ is quasi-closed.*

(iv) *Every subtopos of $\mathcal{E}$ is closed.*

(v) *Every subtopos of $\mathcal{E}$ is open.*

Proof (i) ⇒ (ii): Suppose $\mathcal{E}$ is Boolean; let j be a local operator on $\mathcal{E}$, and A a j-sheaf. We have to show that the lattice of j-closed subobjects of A is Boolean.

But if $A' \rightarrowtail A$ is such a subobject, let $A'' \rightarrowtail A$ be its complement in $\mathrm{Sub}_{\mathcal{E}}(A)$; then $A' \cap c_j(A'') = c_j(A') \cap c_j(A'') = c_j(A' \cap A'') = c_j(0)$ is the least element of $\mathrm{Sub}_{\mathbf{sh}_j(\mathcal{E})}(A)$, and $A' \cup c_j(A'') \geq A' \cup A'' = A$.

(ii) $\Rightarrow$ (iii) follows immediately from 4.5.21.

(iii) $\Rightarrow$ (iv): If (iii) holds, then every closed subtopos is quasi-closed; but this implies that $c(U) = q(U)$ for every subterminal object U. So every subtopos is closed.

(iv) $\Leftrightarrow$ (v): Similarly, if (iv) holds, then every open subtopos is closed; but if $o(U) = c(V)$ then U and V must be complementary elements of Sub(1), and so $c(U) = o(V)$. Hence every closed subtopos is open, so (iv) implies that every subtopos is open. The converse is similar.

(iv) $\Rightarrow$ (i): If (iv) holds, then the double negation operator $\neg\neg$ is closed; but its exterior is 0, so it must be $c(0) = 1_\Omega$. So $\mathcal{E}$ is Boolean by 1.4.14. □

We note that, in a Boolean topos $\mathcal{E}$, a subobject of Ω is classified by a local operator iff it contains $\top$; this may be used to provide an alternative proof of the isomorphism $\mathbf{Lop}(\mathcal{E}) \cong \mathrm{Sub}_{\mathcal{E}}(1)$. (It also shows that the conclusion of 4.5.18 holds whenever the subtopos $\mathbf{sh}_m(\mathcal{E})$ is Boolean; but, in view of 4.5.21 and 4.5.9, this fact may also be deduced from parts (ii) and (iii) of 4.5.18.)

As an application of quasi-closed local operators, we prove an important result, due (in this form) to P. Freyd [371]:

Proposition 4.5.23 *For any topos $\mathcal{E}$, there exists a surjection $f : \mathcal{F} \to \mathcal{E}$ where $\mathcal{F}$ is Boolean.*

Proof We consider the generic subobject $\top : 1 \rightarrowtail \Omega$ as a subterminal object of $\mathcal{E}/\Omega$, and form the topos $\mathcal{F} = \mathbf{sh}_{q(\top)}(\mathcal{E}/\Omega)$. By 4.5.21, $\mathcal{F}$ is Boolean, and it clearly admits a composite geometric morphism to $\mathcal{E}$:

$$\mathcal{F} = \mathbf{sh}_{q(\top)}(\mathcal{E}/\Omega) \longrightarrow \mathcal{E}/\Omega \longrightarrow \mathcal{E}$$

where the second factor is induced by the unique morphism $\Omega \to 1$ in $\mathcal{E}$. We shall show that this morphism is a surjection.

Suppose given a proper monomorphism $m : A' \rightarrowtail A$ in $\mathcal{E}$. We may similarly form the composite geometric morphism

$$f_m : \mathbf{sh}_{q(m)}(\mathcal{E}/A) \longrightarrow \mathcal{E}/A \longrightarrow \mathcal{E}\,.$$

Now in $\mathcal{E}/A$ we have a pullback square

$$\begin{tikzcd} m \arrow[r] \arrow[d, tail] & A^*A' \arrow[d, tail, "A^*m"] \\ 1_A \arrow[r] & A^*A \end{tikzcd}$$

where the bottom edge is the diagonal map $A \to A \times A$. Since both m and 1_A are $q(m)$-sheaves, the left-hand edge of this square remains a proper monomorphism when we apply the associated $q(m)$-sheaf functor to it; and since this functor preserves pullbacks, the same applies to the right-hand edge, i.e. $f_m^*(m)$ is not an isomorphism.

Now let $\phi: \mathcal{E}/A \to \mathcal{E}/\Omega$ be the geometric morphism induced as in 4.1.2 by the classifying map of m. We claim that there is a commutative diagram

$$\begin{tikzcd} \mathbf{sh}_{q(m)}(\mathcal{E}/A) \arrow[r] \arrow[d] & \mathbf{sh}_{q(\top)}(\mathcal{E}/\Omega) \arrow[d] \\ \mathcal{E}/A \arrow[r, "\phi"] & \mathcal{E}/\Omega \end{tikzcd} \quad .$$

To prove this, it suffices by 4.3.12(iii) to show that $\phi_* = \Pi_\phi$ sends m to a $q(\top)$-sheaf, since $q(m)$ is the largest local operator for which m is a sheaf. But $\Pi_\phi(m) \cong \Pi_\phi \phi^*(\top) \cong \top^\phi$ in $\mathcal{E}/\Omega$, so this follows from the fact that sheaves form an exponential ideal (4.4.3). Hence f_m factors through $f_\top = f$; so the fact that $f_m^*(m)$ is not an isomorphism implies that $f^*(m)$ is not an isomorphism. But this is true for any m; so by 1.2.4 f^* is conservative, i.e. f is a surjection. □

We remark that if $\mathcal{E}$ itself is Boolean, then the geometric morphism of 4.5.23 is an equivalence, since in this case we have $\Omega \cong 1 \amalg 1$ and $q(\top) = c(\top) = o(\bot)$ (cf. 4.5.22).

Having considered geometric morphisms from Boolean toposes to a given topos, it is natural to ask what can be said 'in the opposite direction': given a topos $\mathcal{E}$, does there exist a 'nice' geometric morphism $\mathcal{E} \to \mathcal{F}$ where $\mathcal{F}$ is Boolean? Clearly, there can be no such inclusion unless $\mathcal{E}$ itself is Boolean, by 4.5.22(ii); but in general, there may be no such morphism at all. The following argument is due to R. Paré [929]:

Example 4.5.24 Let $\mathcal{E}$ be the topos obtained by glueing along the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$ (cf. 2.1.12). Suppose we have a geometric morphism $f: \mathcal{E} \to \mathcal{F}$ with $\mathcal{F}$ Boolean. By 4.2.10 and 4.5.22(ii), we may assume without loss of generality that f is surjective. Since the objects of $\mathcal{F}$ are all decidable (cf. 1.4.15), every object of the form f^*B is decidable in $\mathcal{E}$; hence, as we observed in 4.2.4(d), it must be an injection between finite sets. It follows that, for any two objects B, B' of $\mathcal{F}$, the set $\mathcal{E}(f^*B, f^*B')$ must be finite, and since f^* is faithful it follows that $\mathcal{F}(B, B')$ is also finite. But, for any set A, let $\hat{A}$ denote the object $(1, A, A \to 1)$ of $\mathcal{E}$; then we have

$$\mathcal{F}(1, f_*\hat{A}) \cong \mathcal{E}(f^*1, \hat{A}) \cong A,$$

so we obtain a contradiction by taking A to be infinite.

A very similar argument (but with the word 'finite', as applied to the size of hom-sets, replaced by 'of cardinality at most that of the continuum') may be used to show that the effective topos **Eff** studied in Chapter F2 also does not admit any geometric morphism to a Boolean topos.

Suggestions for further reading: Barr [71], Borceux & Kelly [151], Freyd [371], Tierney [1169], Wraith [1236].

A4.6 The hyperconnected–localic factorization

In Theorem 4.2.10 we saw that every geometric morphism factors (essentially uniquely) as a surjection followed by an inclusion. There are several other factorization theorems for geometric morphisms, some of which will be discussed in Chapter C3, but one in particular deserves to be mentioned now.

If A and B are objects of a topos, we say A is a *subquotient* of B if there is a diagram of the form

$$\begin{tikzcd} C \arrow[r, two heads] \arrow[d, tail] & A \\ B & \end{tikzcd} ,$$

i.e. A is a quotient of a subobject of B. Since monomorphisms and epimorphisms are stable under pullback and pushout in a topos, this is equivalent to saying that A is a subobject of a quotient of B, i.e. that there is a diagram of the form

$$\begin{tikzcd} & A \arrow[d, tail] \\ B \arrow[r, two heads] & D \end{tikzcd} .$$

It also follows that the relation of being a subquotient is transitive. Yet another way of expressing the condition is to say that there exists a relation $\phi \colon B \nrightarrow A$ (namely, that tabulated by the first displayed diagram above) such that $\phi^\circ \phi$ is the identity relation ι_A, in the notation of Section A3.1. (Such a relation is commonly called a *partial surjection.*)

Definition 4.6.1 We say a geometric morphism $f \colon \mathcal{F} \to \mathcal{E}$ is *localic* if every object of $\mathcal{F}$ is a subquotient of one of the form $f^* A$, where A is an object of $\mathcal{E}$.

The definition is a special case of the notion of a bounded geometric morphism, which we shall meet in B3.1.7; it is equivalent to saying that 1 is a bound

(in the sense defined there) for $\mathcal{F}$ over $\mathcal{E}$. The reason for the name 'localic' will become apparent in C1.4.7.

Examples 4.6.2 (*a*) Every inclusion is localic, for if f is an inclusion then every object of its domain is isomorphic to one of the form f^*A. More generally, if f_* is merely faithful, then the counit $f^*f_*B \to B$ is epic for all B, and so f is localic.

(*b*) Let $f: A \to B$ be a morphism in a topos $\mathcal{E}$. The direct image Π_f of the induced geometric morphism $\mathcal{E}/A \to \mathcal{E}/B$ is not in general faithful; but the left adjoint Σ_f of f^* is, and so the unit $g \to f^*\Sigma_f g$ is monic for all objects g of $\mathcal{E}/A$. Thus all such morphisms are localic.

(*c*) Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories. If f is faithful, then the induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ of 4.1.4 is localic. For every functor $\mathcal{C} \to \mathbf{Set}$ is a quotient of a coproduct of representable functors; if f is faithful then the representable functor $\mathcal{C}(A, -)$ is a subfunctor of $f^*(\mathcal{D}(f(A), -))$; and f^* preserves coproducts. The converse is also true: if $\mathcal{C}(A, -)$ appears as a subquotient of some $f^*(F)$, then (being projective) it actually occurs as a subobject of $f^*(F)$, and this can only happen if there exists $x \in F(f(A))$ such that $F(f\alpha)(x) \neq F(f\beta)(x)$ whenever $\alpha, \beta: A \rightrightarrows B$ are distinct morphisms of $\mathcal{C}$ – which in particular forces $f\alpha \neq f\beta$.

(*d*) In particular, if $\mathcal{C}$ is a preorder (so that the unique functor from $\mathcal{C}$ to the terminal category $\mathbf{1}$ is faithful), then the unique geometric morphism $[\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}$ of 4.1.9 is localic.

(*e*) It is easy to verify that a composite of localic morphisms is localic, since the subquotient relation is transitive and inverse image functors preserve monomorphisms and epimorphisms. So, combining (*a*) and (*d*), we see that if $(\mathcal{C}, T)$ is a small site whose underlying category is a preorder, then the unique geometric morphism $\mathbf{Sh}(\mathcal{C}, T) \to \mathbf{Set}$ is localic. (We shall prove a converse to this result in B3.3.5.) In particular, for any topological space X, $\mathbf{Sh}(X) \to \mathbf{Set}$ is localic. Similarly, combining (*a*) and (*b*), we note that the surjection with Boolean domain constructed in the proof of 4.5.23 is localic.

(*f*) It is even easier to verify that, if

$$\mathcal{G} \xrightarrow{\quad g \quad} \mathcal{F} \xrightarrow{\quad f \quad} \mathcal{E}$$

is a composable pair of geometric morphisms and the composite fg is localic, then g is localic. Hence if $\mathcal{F}$ and $\mathcal{G}$ both admit localic morphisms to $\mathbf{Set}$, then any geometric morphism between them is localic. For example, the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by a continuous map of spaces $X \to Y$, as in 4.1.11, is always localic.

Now let $f: \mathcal{F} \to \mathcal{E}$ be an arbitrary geometric morphism, and let $\mathcal{G}$ be the full subcategory of $\mathcal{F}$ whose objects are all subquotients of objects of the form f^*A.

Lemma 4.6.3 (i) *$\mathcal{G}$ is coreflective in $\mathcal{F}$.*

(ii) *$\mathcal{G}$ is a topos.*

Proof (i) If an object B of $\mathcal{F}$ belongs to $\mathcal{G}$, there is a canonical way of choosing an object A such that B is a subquotient of f^*A, namely $A = f_*\tilde{B}$, where $\tilde{B}$ is the partial-map representer for B (cf. 2.4.7). For if we are given a diagram

$$\begin{tikzcd} C \arrow[r, two heads, "e"] \arrow[d, tail] & B \\ f^*A & \end{tikzcd}$$

there is a unique morphism $a\colon f^*A \to \tilde{B}$ making the appropriate square a pullback; and then a factors through the counit map $f^*f_*\tilde{B} \to \tilde{B}$, which induces a factorization of e through the pullback

$$\begin{tikzcd} P \arrow[r] \arrow[d, tail] & B \arrow[d, tail] \\ f^*f_*\tilde{B} \arrow[r] & \tilde{B} \end{tikzcd}$$

and hence forces the top edge of the above square to be epic.

Now, given an arbitrary object B of $\mathcal{F}$, let us form the pullback P as above, and consider the image factorization $P \twoheadrightarrow GB \rightarrowtail B$. G is clearly a functor $\mathcal{F} \to \mathcal{F}$, since all the constructions involved in its definition are functorial, and it takes values in $\mathcal{G}$ since GB is a subquotient of $f^*f_*\tilde{B}$. Moreover, the monomorphism $GB \rightarrowtail B$ defines a natural transformation from G to the identity, which is an isomorphism precisely when B is in $\mathcal{G}$. It follows immediately from this that GB is the coreflection of B in $\mathcal{G}$.

(ii) It suffices to show that $\mathcal{G}$ is closed under finite limits in $\mathcal{F}$ (equivalently, that the inclusion $\mathcal{G} \to \mathcal{F}$ preserves finite limits), since this implies that the coreflector on $\mathcal{F}$ corresponding to $\mathcal{G}$ is cartesian, and allows us to apply Theorem 4.2.1. But $\mathcal{G}$ is clearly closed under equalizers, since it is closed under arbitrary subobjects, and it contains 1 since f^* preserves 1. Finally, if B_1 and B_2 are subquotients of f^*A_1 and f^*A_2 respectively, then $B_1 \times B_2$ is a subquotient of $f^*(A_1 \times A_2) \cong f^*A_1 \times f^*A_2$. □

We define a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ to be *hyperconnected* if it restricts to an equivalence between $\mathcal{E}$ and the topos $\mathcal{G}$ constructed above; equivalently, if f^* is full and faithful, and its image in $\mathcal{F}$ is closed under arbitrary subobjects and quotients. (The reason for the name 'hyperconnected' will become apparent in C1.5.7, when we introduce the class of connected morphisms.) For example, if G is a topological group, then the morphism $[G, \mathbf{Set}] \to \mathbf{Cont}(G)$ of 4.1.6 is hyperconnected, since every subobject and every quotient of a continuous G-set

is continuous. We shall give further examples of hyperconnected morphisms in 4.6.9 below.

Lemma 4.6.4 *Suppose given a commutative square (up to isomorphism)*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}$, *where* g *is localic and* h *is hyperconnected. Then there exists (uniquely up to unique isomorphism) a morphism* $l\colon \mathcal{F} \to \mathcal{G}$ *making the two triangles commute.*

Proof Let B be an object of $\mathcal{G}$. Then B is a subquotient of g^*A for some A, whence k^*B is a subquotient of $k^*g^*A \cong h^*f^*A$. Since the image of h^* is closed under subquotients, it follows that k^*B lies in it, and so (since h^* is full and faithful) we get a unique factorization $k^* = h^*l^*$, where l^* is a functor $\mathcal{G} \to \mathcal{F}$. As in the proof of 4.3.11, we now see that $l_* = k_*h^*$ is a right adjoint for l^*; moreover, l^* preserves finite limits (because h^* creates them), and the isomorphism $h^*l^*g^* = k^*g^* \cong h^*f^*$ yields an isomorphism $l^*g^* \cong f^*$ because h^* is full and faithful. This establishes the existence of l; the uniqueness again follows from the fullness and faithfulness of h^*. □

Theorem 4.6.5 *Any geometric morphism can be factored, uniquely up to equivalence, as a hyperconnected morphism followed by a localic one.*

Proof The existence of the factorization was mostly established in 4.6.3: given a morphism $f\colon \mathcal{F} \to \mathcal{E}$, we let $\mathcal{G}$ be the topos defined there, $h\colon \mathcal{F} \to \mathcal{G}$ the geometric morphism whose inverse image is the inclusion $\mathcal{G} \to \mathcal{F}$, and $g\colon \mathcal{G} \to \mathcal{E}$ the morphism whose inverse image is the factorization of f^* through this inclusion (its direct image, once again, being the composite f_*h^*). Then it is easily seen that h is hyperconnected, that g is localic and that $f \cong gh$. The uniqueness follows straightforwardly from 4.6.4. □

There are a number of alternative characterizations of hyperconnected morphisms which are of interest; we now list them.

Proposition 4.6.6 *Let* $f\colon \mathcal{F} \to \mathcal{E}$ *be a geometric morphism. The following are equivalent:*

(i) f *is hyperconnected.*

(ii) f^* *is full and faithful, and its image is closed under subobjects in* $\mathcal{F}$.

(iii) f^* *is full and faithful, and its image is closed under quotients in* $\mathcal{F}$.

(iv) *The unit and counit of* $(f^* \dashv f_*)$ *are both monic.*

(v) f_* *preserves* Ω*, i.e. the comparison map* $\tau \colon f_*(\Omega_{\mathcal{F}}) \to \Omega_{\mathcal{E}}$ *(the classifying map of* $f_*(\top_{\mathcal{F}})$*) is an isomorphism.*

(vi) *For each object* A *of* $\mathcal{E}$*,* f^* *induces an equivalence* $\mathrm{Sub}_{\mathcal{E}}(A) \simeq \mathrm{Sub}_{\mathcal{F}}(f^*A)$.

Proof (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are trivial.

(iii) $\Rightarrow$ (iv): If f^* is full and faithful, then the unit of $(f^* \dashv f_*)$ is an isomorphism, and hence monic. Now consider the image factorization $f^*f_*B \twoheadrightarrow I \rightarrowtail B$ of the counit ϵ_B: by assumption I is in the image of f^*, so by the universal property of ϵ_B the epic part of this factorization must be an isomorphism, i.e. ϵ_B is monic.

(iv) $\Rightarrow$ (iii): Let η and ϵ be the unit and counit of $(f^* \dashv f_*)$. From one of the triangular identities and the fact that ϵ is monic, we deduce that $f^*\eta$ is an isomorphism; but since η is monic f^* is conservative (cf. 4.2.6), and so η is an isomorphism, i.e. f^* is full and faithful. Now we may identify $\mathcal{E}$ (up to equivalence) with a coreflective subcategory of $\mathcal{F}$; if B is a quotient of an object in this subcategory, then we see that the counit $\epsilon_B \colon f^*f_*B \to B$ must be epic as well as monic, and so B is in the subcategory.

(iii) $\Rightarrow$ (v): Let us identify $\mathcal{E}$ with a coreflective subcategory of $\mathcal{F}$, and write G for the coreflector f^*f_*. Then the square

$$\begin{tikzcd}
G(1) \arrow[r] \arrow[d, tail, "G(\top)"] & 1 \arrow[d, tail, "\top"] \\
G(\Omega) \arrow[r, "\epsilon_\Omega"] & \Omega
\end{tikzcd}$$

commutes by naturality of ϵ, and since ϵ_Ω is monic by (iv) (and G preserves 1) it must be a pullback. Now if $A' \rightarrowtail A$ is a subobject in $\mathcal{E}$, the classifying map $\phi \colon A \to \Omega$ of A' in $\mathcal{F}$ factors uniquely through ϵ_Ω; and in the diagram

$$\begin{tikzcd}
A' \arrow[r] \arrow[d, tail] & G(1) \arrow[r] \arrow[d, tail, "G(\top)"] & 1 \arrow[d, tail, "\top"] \\
A \arrow[r, "\bar{\phi}"] & G(\Omega) \arrow[r, "\epsilon_\Omega"] & \Omega
\end{tikzcd}$$

the outer and right-hand squares are pullbacks, whence the left-hand square is a pullback. But the requirement that this square be a pullback determines $\epsilon_\Omega\bar{\phi}$, and hence $\bar{\phi}$, uniquely; thus $G(\Omega)$ is a subobject classifier for $\mathcal{E}$.

(ii) $\Rightarrow$ (v): Again, we identify $\mathcal{E}$ with a coreflective subcategory of $\mathcal{F}$, and write G for the coreflector. This time we form the pullback

$$\begin{tikzcd} P \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\ G(\Omega) \arrow[r, "\epsilon_\Omega"] & \Omega \end{tikzcd}$$

and observe that P, being a subobject of $G(\Omega)$, belongs to $\mathcal{E}$. As in the previous part, we see that every subobject in $\mathcal{E}$ is uniquely expressible as a pullback of $P \rightarrowtail G(\Omega)$, i.e. the latter is a generic subobject in $\mathcal{E}$. So by 1.6.1 P is a terminal object of $\mathcal{E}$, and hence of $\mathcal{F}$; thus $P \rightarrowtail G(\Omega)$ is isomorphic to $G(\top)$ in $\mathrm{Sub}(G(\Omega))$ (since the latter factors through the former, and both have domain 1). So $G(\top)$ is a generic subobject in $\mathcal{E}$, i.e. G preserves Ω.

(v) $\Rightarrow$ (vi): It is immediate from (v) (and the adjunction $(f^* \dashv f_*)$) that we have a bijection between (isomorphism classes of) subobjects of A and of f^*A, for any object A of $\mathcal{E}$. However, as in 2.4.8, it requires proof that this bijection is induced by applying f^* to monomorphisms in $\mathcal{E}$. (Unfortunately, we cannot employ the argument in the proof of 2.4.8, since we do not know that f_* preserves partial-map representers.) Let $\tau' \colon f^*\Omega_{\mathcal{E}} \to \Omega_{\mathcal{F}}$ be the comparison map for f^* (i.e. the classifying map of $f^*(\top_{\mathcal{E}})$); then the bijection induced by the fact that f_* preserves Ω sends the subobject classified by $\phi \colon A \to \Omega_{\mathcal{E}}$ to the subobject classified by the composite

$$f^*A \xrightarrow{f^*\phi} f^*\Omega_{\mathcal{E}} \xrightarrow{f^*\tau^{-1}} f^*f_*\Omega_{\mathcal{F}} \xrightarrow{\epsilon_\Omega} \Omega_{\mathcal{F}},$$

whereas its image under f^* is the subobject classified by

$$f^*A \xrightarrow{f^*\phi} f^*\Omega_{\mathcal{E}} \xrightarrow{\tau'} \Omega_{\mathcal{F}}\,.$$

So we need to show that the diagram

$$\begin{tikzcd} f^*f_*\Omega_{\mathcal{F}} \arrow[r, "f^*\tau"] \arrow[dr, "\epsilon_\Omega"'] & f^*\Omega_{\mathcal{E}} \arrow[d, "\tau'"] \\ & \Omega_{\mathcal{F}} \end{tikzcd}$$

commutes. Since the square

$$\begin{tikzcd} f^*f_*1 \arrow[r, "\epsilon_1"] \arrow[d, tail, "f^*f_*\top"'] & 1 \arrow[d, tail, "\top"] \\ f^*f_*\Omega_{\mathcal{F}} \arrow[r, "\epsilon_\Omega"] & \Omega_{\mathcal{F}} \end{tikzcd}$$

commutes (though we don't yet know it is a pullback), the subobject of $f^*f_*\Omega_{\mathcal{F}}$ classified by ϵ_Ω contains that classified by $\tau'(f^*\tau)$, and hence the subobject of f^*A corresponding under the bijection to $A' \rightarrowtail A$ contains f^*A'. Thus f^* preserves properness of subobjects; since it also preserves equalizers, it is conservative by 1.2.4, and so the unit η of $(f^* \dashv f_*)$ is monic. Now, in the diagram

$$\begin{tikzcd} 1_{\mathcal{E}} \arrow[r, tail, "\eta_1"] \arrow[d, tail, "\top_{\mathcal{E}}"'] & f_*f^*1_{\mathcal{E}} \arrow[r] \arrow[d, tail, "f_*f^*\top_{\mathcal{E}}"'] & f_*1_{\mathcal{F}} \arrow[r] \arrow[d, tail, "f_*\top_{\mathcal{F}}"'] & 1_{\mathcal{E}} \arrow[d, tail, "\top_{\mathcal{E}}"'] \\ \Omega_{\mathcal{E}} \arrow[r, tail, "\eta_\Omega"] & f_*f^*\Omega_{\mathcal{E}} \arrow[r, "f_*\tau'"] & f_*\Omega_{\mathcal{F}} \arrow[r, "\tau"] & \Omega_{\mathcal{E}} \end{tikzcd}$$

all three squares are pullbacks, so the bottom composite is the identity, and since τ is an isomorphism the composite $(f_*\tau')\eta_\Omega\tau$ is the identity on $f_*\Omega_{\mathcal{F}}$. Transposing this composite across the adjunction $(f^* \dashv f_*)$, we obtain $\tau'(f^*\tau)$, so the latter must equal ϵ_Ω, as required.

(vi) $\Rightarrow$ (i): First we note that (vi) implies that f^* preserves properness of subobjects, so it is conservative by 1.2.4. Next, let $a\colon f^*A \to f^*B$ be a morphism in $\mathcal{F}$ between objects in the image of f^*. Now the graph $(1,a)\colon f^*A \rightarrowtail f^*A \times f^*B \cong f^*(A \times B)$ is isomorphic to the image under f^* of some subobject $(p,q)\colon A' \rightarrowtail A \times B$; but then f^*p is an isomorphism and so p is an isomorphism, i.e. (p,q) is the graph of a morphism $qp^{-1}\colon A \to B$, whose image under f^* is clearly a. So f^* is full as well as faithful. It is immediate from (vi) that the image of f^* is closed under subobjects in $\mathcal{F}$, so it remains to verify that it is closed under quotients. Given an epimorphism $f^*A \twoheadrightarrow B$, let $S \rightarrowtail f^*A \times f^*A \cong f^*(A \times A)$ be its kernel-pair; then S must be of the form f^*R for some $R \rightarrowtail A \times A$. Let $A \twoheadrightarrow A'$ be the coequalizer of $R \rightrightarrows A$; then since f^* preserves coequalizers we have $f^*A' \cong B$. □

Remark 4.6.7 Combining 4.6.6(v) with 4.2.9, we obtain another proof of the result (already established in 2.3.9) that a geometric morphism whose direct image is logical must be an equivalence.

We digress briefly to note a connection between hyperconnected morphisms and the weak Beck–Chevalley condition introduced in 4.1.16.

Lemma 4.6.8 *Suppose given a commutative square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"'] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}$, *such that* either f *and* k *are both hyperconnected,* or g *and* h *are both hyperconnected. Then the square satisfies the weak Beck–Chevalley condition, i.e. the canonical natural transformation* $\theta\colon g^* f_* \to k_* h^*$ *is pointwise monic.*

Proof First suppose f and k are hyperconnected. As we saw in the proof of 4.6.6, this implies that f^* and f_* restrict to an equivalence between the categories of subterminal objects in $\mathcal{E}$ and in $\mathcal{F}$; and similarly for k^* and k_*. Hence, for any subterminal object B of $\mathcal{F}$, we have $g^* f_* B \cong k_* k^* g^* f_* B \cong k_* h^* f^* f_* B \cong k_* h^* B$, so θ_B must be an isomorphism. Since the property of hyperconnectedness is 'stable under slicing' in the obvious sense, it follows that the square satisfies condition (iii) of 4.1.17. The case when g and h are hyperconnected is proved similarly, using the natural isomorphism $f_* h_* \cong g_* k_*$. □

Next, we re-examine Example 4.6.2(*c*) in the light of the results established in this section.

Example 4.6.9 Let $f\colon \mathcal{C} \to \mathcal{D}$ be a functor between small categories. Suppose f is full and bijective on objects; then $\mathcal{D}$ may be identified with the quotient category $\mathcal{C}/R$, where R is the congruence induced by f (i.e. the equivalence relation on morphisms such that $(\alpha, \beta) \in R$ iff $f\alpha = f\beta$), and $[\mathcal{D}, \mathbf{Set}]$ is identified via f^* with the full subcategory of $[\mathcal{C}, \mathbf{Set}]$ consisting of functors that respect this congruence (equivalently, factor through f). Since this subcategory is clearly closed under subobjects and quotients, it follows that the geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ induced by f is hyperconnected. Combining this with 4.6.2(*c*), we see that for an arbitrary f the factorization of 4.6.5, applied to the geometric morphism induced by f, corresponds to the unique factorization $\mathcal{C} \to \mathcal{C}/R \to \mathcal{D}$ of f as a full functor bijective on objects followed by a faithful functor. (This result should be compared with 4.2.12(*b*).)

Restricting this example to monoids (small categories with one object), we see that in this context the factorization of 4.6.5 corresponds to the usual cover–monic factorization in the (regular) category of monoids. In contrast, the factorization of 4.2.10 yields nothing at all here, since the geometric morphism induced by a monoid homomorphism is always surjective. (On the other hand, for the geometric morphisms between sheaf toposes induced by continuous maps of spaces, it is the factorization of 4.6.5 that yields nothing, since all such morphisms are localic.) Note also that it follows from 4.6.9 that $[\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}$ is

hyperconnected iff $\mathcal{C}$ is *strongly connected*, i.e., for any two objects A and B of $\mathcal{C}$, there exist morphisms $A \to B$ and $B \to A$.

We have observed that every inclusion is localic, and similarly every hyperconnected morphism is surjective. It follows easily that the two factorizations of 4.2.10 and 4.6.5 coincide (up to equivalence), for a given f, precisely when f can be factored as a hyperconnected morphism followed by an inclusion. In this connection, we have

Proposition 4.6.10 *For a geometric morphism $f: \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *The surjection–inclusion and hyperconnected–localic factorizations of f coincide.*

(ii) *f can be factored as a composite of morphisms which are either hyperconnected or inclusions.*

(iii) *The counit of the adjunction $(f^* \dashv f_*)$ is monic.*

Proof (i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii) because both hyperconnected morphisms and inclusions satisfy (iii), and it is easily verified that this property is stable under composition.

(iii) $\Rightarrow$ (i): Let $f \cong gh$ be the surjection–inclusion factorization of f. By the construction of 4.2.8, the counit of $(h^* \dashv h_*)$ coincides with that of $(f^* \dashv f_*)$ and is therefore monic; but the unit of $(h^* \dashv h_*)$ is also monic by 4.2.6(iv), and so h is hyperconnected by 4.6.6(iv). □

For a functor $f: \mathcal{C} \to \mathcal{D}$ between small categories, it is easy to see that the induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ satisfies the conditions of 4.6.10 iff f is full.

A further result relating hyperconnected morphisms and inclusions, which will be of importance later on, is the following.

Proposition 4.6.11 *Let $f: \mathcal{F} \to \mathcal{E}$ be a hyperconnected morphism, let j be a local operator on $\mathcal{E}$, and let k be the smallest local operator on $\mathcal{F}$ for which the composite $\mathbf{sh}_k(\mathcal{F}) \to \mathcal{F} \to \mathcal{E}$ factors through $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$ (cf. 4.5.14(e)). Then the resulting factorization $f': \mathbf{sh}_k(\mathcal{F}) \to \mathbf{sh}_j(\mathcal{E})$ is hyperconnected.*

Proof First we observe that since f^* induces an isomorphism $\mathrm{Sub}_{\mathcal{E}}(A) \to \mathrm{Sub}_{\mathcal{F}}(f^*A)$ for any A, it preserves the Heyting implication; hence if $A' \rightarrowtail A$ is a j-dense subobject and $A'' \rightarrowtail A$ is j-closed, then the pair (f^*A', f^*A'') must satisfy the relation θ defined before 4.5.12. In particular, the composite

$$f^*J \times f^*\Omega_j \rightarrowtail f^*\Omega_{\mathcal{E}} \times f^*\Omega_{\mathcal{E}} \xrightarrow{\tau' \times \tau'} \Omega_{\mathcal{F}} \times \Omega_{\mathcal{F}}$$

factors through $\Theta_{\mathcal{F}} \rightarrowtail \Omega_{\mathcal{F}} \times \Omega_{\mathcal{F}}$ (note, incidentally, that the comparison map τ' is monic, by the proof of (v) $\Rightarrow$ (vi) in 4.6.6). Since $K \rightarrowtail \Omega_{\mathcal{F}}$ is defined to be $(f^*J \rightarrowtail f^*\Omega_{\mathcal{E}} \rightarrowtail \Omega_{\mathcal{F}})^{rl}$, it follows that we have $f^*\Omega_j \leq (f^*J)^r = K^r = \Omega_k$ in

$\mathrm{Sub}(\Omega_{\mathcal{F}})$, and hence that f^* maps j-closed monomorphisms to k-closed ones. In particular, if A is a j-sheaf in $\mathcal{E}$, then f^*A is k-separated by 4.3.6(a)(iv); hence the unit map $A \to f'_* f'^* A$ is monic, since it is obtained by applying f_* to the unit map from f^*A to its associated k-sheaf.

Thus we have shown that f' is a surjection; hence $\mathbf{sh}_j(\mathcal{E})$ is the image (in the surjection–inclusion sense) of the composite $\mathbf{sh}_k(\mathcal{F}) \to \mathcal{F} \to \mathcal{E}$. The result is now immediate from 4.6.10, since the composite just mentioned satisies condition (ii) of 4.6.10. □

The 'orthogonality' result 4.6.4 tells us that, for a fixed topos $\mathcal{E}$, the hyperconnected–localic factorization applied to geometric morphisms with codomain $\mathcal{E}$ yields a left adjoint to the inclusion $\mathfrak{LTop}/\mathcal{E} \to \mathfrak{Top}/\mathcal{E}$, where $\mathfrak{LTop}$ is the 2-category of toposes, localic geometric morphisms and geometric transformations between them. (Note that $\mathfrak{LTop}/\mathcal{E}$ is a full sub-2-category of $\mathfrak{Top}/\mathcal{E}$, by 4.6.2($f$).) We refer to this left adjoint as the *localic reflection* for toposes over $\mathcal{E}$.

Proposition 4.6.12 *The localic reflection functor $\mathfrak{Top}/\mathcal{E} \to \mathfrak{LTop}/\mathcal{E}$ preserves surjections and inclusions.*

Proof Let $f_i \colon \mathcal{F}_i \to \mathcal{E}$ $(i = 1, 2)$ be two objects of $\mathfrak{Top}/\mathcal{E}$, and let $g \colon \mathcal{F}_1 \to \mathcal{F}_2$ be a geometric morphism over $\mathcal{E}$. Form the diagram

$$\begin{tikzcd} \mathcal{F}_1 \arrow[r, "g"] \arrow[d, "h_1"] & \mathcal{F}_2 \arrow[d, "h_2"] \\ \mathcal{L}_1 \arrow[r, "k"] \arrow[dr, "l_1"'] & \mathcal{L}_2 \arrow[d, "l_2"] \\ & \mathcal{E} \end{tikzcd}$$

where (h_i, l_i) is the hyperconnected–localic factorization of f_i and k is the morphism obtained by applying the orthogonality result 4.6.4 to h_1 and l_2. It is clear that if g is surjective, then so is the composite $h_2 g \cong k h_1$, and hence k is surjective. Suppose g is an inclusion; by 4.6.2(f), we know that k is localic, and hence the left and bottom edges of the square form the hyperconnected–localic factorization of its diagonal. But since g is an inclusion and h_2 is hyperconnected, it follows from 4.6.10 that this must also be the surjection–inclusion factorization of the diagonal; so k is an inclusion. □

Further preservation properties of the localic reflection will be encountered in C2.4.13 and C2.5.12.

Suggestions for further reading: Johnstone [518], Joyal & Tierney [560].

PART B

2-CATEGORICAL ASPECTS OF TOPOS THEORY

B1

INDEXED CATEGORIES AND FIBRATIONS

B1.1 Review of 2-categories

For many of the topics we wish to study in topos theory, it is impossible to ignore the fact that toposes and geometric morphisms form not just a category but a 2-category; that is, we have to take account also of the geometric transformations between geometric morphisms, as introduced in A4.1.1(b). Even amongst experienced category-theorists, there is often a good deal of reluctance to take 2-categorical (or, still worse, bicategorical) notions seriously; so the purpose of this section is to review as much of the basic theory as we shall need for subsequent developments.

The neatest definition of a 2-category is as a **Cat**-enriched category, where **Cat** denotes the category of small categories. In other words, a 2-category $\mathfrak{K}$ consists of a class of objects $A, B, \ldots$ together with, for each pair of objects (A, B), a category $\mathfrak{K}(A, B)$, and for each triple (A, B, C) a composition operation

$$\mathfrak{K}(B,C) \times \mathfrak{K}(A,B) \longrightarrow \mathfrak{K}(A,C)$$

which is bifunctorial and associative, and equipped with identities. However, since this definition begs the question of local smallness (and because we have not yet reviewed the theory of enriched categories – see Section B2.1 below), we shall spell out in full what this means in elementary terms.

Definition 1.1.1 A *2-category* $\mathfrak{K}$ consists of three types of entities: *objects* or *0-cells* (usually denoted $A, B, C, \ldots$), *morphisms* or *1-cells* (denoted $f, g, h, \ldots$), and *transformations* or *2-cells* (denoted $\alpha, \beta, \gamma, \ldots$), equipped with the following structure:

(a) There are two operations dom, cod from morphisms to objects.

(b) There are two operations sce, tgt from transformations to morphisms, satisfying dom(sce α) = dom(tgt α) and cod(sce α) = cod(tgt α) for all α. (We shall abbreviate dom(sce α) to dom α and cod(sce α) to cod α when this is unlikely to cause confusion.)

(c) There is an operation $A \mapsto 1_A$ from objects to morphisms, which is a simultaneous splitting for dom and cod.

(*d*) There is an operation $f \mapsto \iota_f$ from morphisms to transformations, which is a simultaneous splitting for sce and tgt.

(*e*) There is a partial binary operation $(f, g) \mapsto gf$ on morphisms, such that gf is defined iff cod f = dom g, and the equations dom gf = dom f, cod gf = cod g, $h(gf) = (hg)f$, $f1_A = f$ and $1_B f = f$ hold whenever they make sense.

(*f*) There are two partial binary operations on transformations: vertical composition (denoted $(\alpha, \beta) \mapsto \beta\alpha$) and horizontal composition (denoted $(\alpha, \beta) \mapsto \beta \circ \alpha$), such that

  (i) $\beta\alpha$ is defined iff tgt α = sce β and $\beta \circ \alpha$ is defined iff cod α = dom β;

  (ii) sce $\beta\alpha$ = sce α and tgt $\beta\alpha$ = tgt β;

  (iii) sce $\beta \circ \alpha$ = (sce β)(sce α) and tgt $\beta \circ \alpha$ = (tgt β)(tgt α);

  (iv) the associative laws $\gamma(\beta\alpha) = (\gamma\beta)\alpha$ and $\gamma \circ (\beta \circ \alpha) = (\gamma \circ \beta) \circ \alpha$ hold;

  (v) transformations of the form ι_f act as units for vertical composition, and those of the form ι_{1_A} act as units for horizontal composition, and

  (vi) the interchange law $(\delta\gamma) \circ (\beta\alpha) = (\delta \circ \beta)(\gamma \circ \alpha)$ and the identity $\iota_g \circ \iota_f = \iota_{gf}$ hold,

it being understood that all the above identities are to be read in the sense 'if both sides are defined, then they are equal'.

The names 'horizontal composition' and 'vertical composition' are suggested by the pictures

$$\begin{tikzcd} A \arrow[r, shift left=1.5ex, "f"] \arrow[r, shift right=1.5ex, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & B \arrow[r, shift left=1.5ex, "h"] \arrow[r, shift right=1.5ex, "k"'] \arrow[r, phantom, "\downarrow\beta"] & C \end{tikzcd}$$

and

$$\begin{tikzcd} A \arrow[r, shift left=3ex, "f"] \arrow[r, "g" description] \arrow[r, shift left=1.5ex, phantom, "\downarrow\alpha"] \arrow[r, shift right=1.5ex, phantom, "\downarrow\beta"] \arrow[r, shift right=3ex, "h"'] & B \end{tikzcd}$$

which we frequently draw to provide a visual description of what is going on inside a 2-category. We shall not go into the detailed theory of such 'pasting diagrams' here, but we shall feel free to use it at least in an informal way. In passing, we remark that if one of the 2-cells in a horizontal composite is a (vertical) identity ι_f, then we normally write $\alpha \circ f$ instead of $\alpha \circ \iota_f$.

Of course, if A and B are two objects of a 2-category $\mathfrak{K}$, then the morphisms and transformations of $\mathfrak{K}$ with domain A and codomain B form a category, which we denote by $\mathfrak{K}(A, B)$. Composition with a fixed morphism of $\mathfrak{K}$ induces functors between these categories, and horizontal composition with a fixed transformation induces natural transformations between such functors. It should be

noted that we have not imposed any smallness condition (not even local smallness) on the hom-categories of $\mathfrak{K}$, and in many of the examples we shall have to consider this lack of restriction will be essential. However, in introducing 2-categorical concepts we shall often 'cheat' by pretending that our 2-categories do have 'category-valued hom-functors', and the like, simply in order to avoid the tedium of writing out in full the elementary version of what the definitions mean. (But the reader should be aware that it is always possible to do so, if one is sufficiently determined!) In this connection, we shall use the notation $\mathfrak{CAT}$ for the 'meta-2-category' of categories (without size restriction), functors and natural transformations, as opposed to the (legitimate) 2-category $\mathfrak{Cat}$ of small categories.

There are two different duality principles that we can apply to an arbitrary 2-category $\mathfrak{K}$, one involving the reversal of 1-cells and the other the reversal of 2-cells. It is convenient to distinguish notationally between these; accordingly, if $\mathfrak{K}$ is a 2-category, we shall write $\mathfrak{K}^{\mathrm{op}}$ for the 2-category obtained by reversing the 1-cells but not the 2-cells of $\mathfrak{K}$ (so that $\mathfrak{K}^{\mathrm{op}}(A,B) = \mathfrak{K}(B,A)$), and $\mathfrak{K}^{\mathrm{co}}$ for that obtained by reversing 2-cells but not 1-cells (so that $\mathfrak{K}^{\mathrm{co}}(A,B) = (\mathfrak{K}(A,B))^{\mathrm{op}}$). If we need to reverse both the 1-cells and the 2-cells of $\mathfrak{K}$, we denote the result by $\mathfrak{K}^{\mathrm{coop}}$.

If P is a property of categories, we shall say that a 2-category is *locally* P if its hom-categories all have property P; thus a locally small 2-category is precisely a **Cat**-enriched category. This usage conflicts with our normal use of the word 'locally' for ordinary categories, as introduced in Section A1.1 (though not with our use of 'locally small' for ordinary categories!), but it seems unlikely to lead to confusion. In particular, if $\mathcal{C}$ is an ordinary category, we shall identify it when necessary with the *locally discrete* 2-category having the same objects and morphisms as $\mathcal{C}$, but only (vertical) identity transformations.

A guiding principle of our approach to 2-category theory (which, incidentally, is part of the reason for not regarding 2-categories simply as **Cat**-enriched categories) is that it is in most cases unreasonable to expect diagrams in a given 2-category to commute 'on the nose'; usually we are much more interested in diagrams which commute up to 2-isomorphism (that is, up to invertible 2-cells placed in every cell of the diagram). These 2-isomorphisms must of course satisfy 'coherence conditions' which ensure that any two composite 2-cells constructed by pasting them together, and having the same source and target, are equal. In the interests of maintaining clarity, we shall normally not spell out such conditions in their fullest detail; indeed, we shall often suppress the names of the 'canonical' 2-isomorphisms up to which our diagrams commute, and simply write $f \cong g$ (where f and g are the names of composite 1-cells) where in an ordinary category we would write $f = g$, to indicate that we are considering the canonical 2-isomorphism between them. We shall also occasionally use the notation $\alpha \approx \beta$, where α and β are 2-cells whose sources and targets are not equal but only canonically isomorphic, to indicate that they become equal on composition with the canonical isomorphisms. But it is important to remember that the

canonical 2-isomorphisms really are there, and must be carried along as part of the structure we are talking about.

In particular, when we speak of a *functor* between 2-categories, we generally mean what is officially called a pseudofunctor. (We shall also sometimes call it a *2-functor*, to emphasize the fact that it acts on 2-cells as well as objects and 1-cells of the domain category.)

Definition 1.1.2 Let $\mathfrak{K}$ and $\mathfrak{L}$ be 2-categories. A *(pseudo)functor* $F\colon \mathfrak{K} \to \mathfrak{L}$ consists of the following data:

(*a*) a mapping $A \mapsto FA$ from objects of $\mathfrak{K}$ to objects of $\mathfrak{L}$;

(*b*) a mapping $f \mapsto Ff$ from morphisms of $\mathfrak{K}$ to morphisms of $\mathfrak{L}$;

(*c*) a mapping $\alpha \mapsto F\alpha$ from transformations of $\mathfrak{K}$ to transformations of $\mathfrak{L}$;

(*d*) for each object A of $\mathfrak{K}$, a 2-isomorphism $\phi_A\colon 1_{FA} \to F(1_A)$, and

(*e*) for each composable pair (f, g) of morphisms of $\mathfrak{K}$, a 2-isomorphism $\phi_{f,g}\colon (Fg)(Ff) \to F(gf)$,

such that dom Ff = F(dom f), cod Ff = F(cod f), sce $F\alpha$ = F(sce α), tgt $F\alpha$ = F(tgt α), $F(\iota_f) = \iota_{Ff}$, $F(\beta\alpha) = (F\beta)(F\alpha)$, and the ϕ's satisfy the coherence conditions that

$$\begin{tikzcd}
FA \arrow[r, bend left, "1_{FA}"] \arrow[r, phantom, "\downarrow \phi_A"] \arrow[r, "F(1_A)"'] \arrow[dr, bend right=60, "Ff"'] \arrow[dr, phantom, "\downarrow \phi_{1_A,f}"] & FA \arrow[d, "Ff"] \\
& FB
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
FA \arrow[r, "Ff"] \arrow[dr, bend right=60, "Ff"'] \arrow[dr, phantom, "\overleftarrow{\phi_{f,1_B}}"] & FB \arrow[d, "F(1_B)"'] \arrow[d, bend left, "1_{FB}"] \arrow[d, phantom, "\overleftarrow{\phi_B}"] \\
& FB
\end{tikzcd}$$

are both equal to ι_{Ff}, that the two composites

$$\begin{tikzcd}
FB \arrow[r, "Fg"] \arrow[r, phantom, "\downarrow \phi_{h,g}", shift right=4] & FC \arrow[d, "Ff"] \\
FA \arrow[u, "Fh"] \arrow[ur, "F(gh)"'] \arrow[r, "F(fgh)"'] \arrow[ur, phantom, "\downarrow \phi_{gh,f}", shift right=6] & FD
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
FB \arrow[r, "Fg"] \arrow[dr, "F(fg)"'] \arrow[dr, phantom, "\downarrow \phi_{g,f}", shift left=6] & FC \arrow[d, "Ff"] \\
FA \arrow[u, "Fh"] \arrow[r, "F(fgh)"'] \arrow[ur, phantom, "\downarrow \phi_{h,fg}", shift right=6] & FD
\end{tikzcd}$$

are equal, and that

$$\phi_{g,k}(F\beta \circ F\alpha) = F(\beta \circ \alpha)\phi_{f,h}$$

for any horizontally composable pair of 2-cells

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & B \arrow[r, bend left, "h"] \arrow[r, bend right, "k"'] \arrow[r, phantom, "\downarrow\beta"] & C . \end{tikzcd}$$

In practice, the pseudofunctors we meet are generally *normalized* (or *strictly unital*); that is, they satisfy $F(1_A) = 1_{FA}$ and the ϕ_A are identity morphisms. We shall often assume this condition without commenting on it – after all, if we are given a pseudofunctor which happens not to be strictly unital, we can easily 'normalize' it by redefining $F(1_A)$ to be 1_{FA} for all A (and suitably modifying the definition of $\phi_{f,g}$ whenever either f or g is an identity morphism). However, there is no similar way of modifying a given F so that the $\phi_{f,g}$ all become identities, and so we shall not assume the latter condition without saying so explicitly. (We say so by calling F a *strict functor.*)

Occasionally we shall wish to consider *lax functors*, which differ from pseudofunctors in that the ϕ_A and $\phi_{f,g}$ are not even required to be isomorphisms; and *oplax functors*, which are similar except that the ϕ_A and $\phi_{f,g}$ all point in the opposite direction. (In other words, an oplax functor $\mathfrak{K} \to \mathfrak{L}$ is the same thing as a lax functor $\mathfrak{K}^{\mathrm{co}} \to \mathfrak{L}^{\mathrm{co}}$.)

There is, perhaps, a certain lack of consistency in the fact that we have standardized on pseudofunctors rather than strict 2-functors as the basic notion, but also on 'strict' 2-categories rather than *bicategories*, in which the associative and unit laws for composition of morphisms hold only up to (specified, coherent) 2-isomorphisms. Our justification for this is basically pragmatic: virtually all the 2-dimensional categories we shall meet (with one important exception, the bicategory of profunctors which we shall encounter in Section B2.7) are in fact strict 2-categories (usually this is because their morphisms are in fact functions of some kind, and their composition is the usual composition of functions), and it saves notational complications if we treat them as such. Nevertheless, nearly all the results about 2-categories which we review here are also true for bicategories; we refer the reader to [99] for further details.

Between (pseudo)functors, we have *(pseudo)natural transformations*; such a transformation between functors (F, ϕ) and $(G, \psi)\colon \mathfrak{K} \to \mathfrak{L}$ assigns to each object A of $\mathfrak{K}$ a morphism $e_A\colon FA \to GA$ of $\mathfrak{L}$, and to each morphism $f\colon A \to B$ of $\mathfrak{K}$ a 2-isomorphism η_f fitting into the diagram

$$\begin{tikzcd} FA \arrow[r, "Ff"] \arrow[d, "e_A"] \arrow[dr, phantom, "\downarrow\eta_f"] & FB \arrow[d, "e_B"] \\ GA \arrow[r, "Gf"] & GB \end{tikzcd}$$

subject to the evident compatibility conditions with the ϕ's and ψ's. As with functors, we shall occasionally want to consider *lax natural transformations*, which are defined in exactly the same way but without the requirement that the η_f should be isomorphisms.

There is also a fourth level of structure, that of *modifications* between natural transformations: a modification between natural transformations (e, η) and (e', η') assigns to each object A of $\mathfrak{K}$ a transformation $e_A \to e'_A$ in $\mathfrak{L}$, subject to compatibility conditions with the η_f and η'_f.

In any 2-category, we say a morphism $f: A \to B$ is *left adjoint* to $g: B \to A$ (and write $(f \dashv g)$) if there are transformations $\eta: 1_A \to gf$ and $\epsilon: fg \to 1_B$ (the *unit* and *counit* of the adjunction) satisfying the usual triangular identities. The right adjoint of a given morphism, if it exists, is unique up to unique 2-isomorphism. An *equivalence* in a 2-category may conveniently be defined as an adjunction whose unit and counit are 2-isomorphisms; we shall also use the terms *reflection* and *coreflection* for adjunctions where just one of these transformations is invertible.

Similarly, given functors F and $G: \mathfrak{K} \to \mathfrak{L}$, a *natural adjunction* between F and G consists of a pair of natural transformations $d: F \to G$, $e: G \to F$, equipped with modifications $\eta: 1 \to ed$ and $\epsilon: de \to 1$ satisfying the triangular identities; what this means is that we have, for each object A of $\mathfrak{K}$, an adjunction $(d_A \dashv e_A)$ in $\mathfrak{L}$, and that these adjunctions 'fit together' – that is, they satisfy appropriate compatibility conditions with the coherence 2-isomorphisms (whose names we have already suppressed) forming part of the specification of d and e. We shall be mainly concerned with *natural equivalences*, i.e. natural adjunctions which are 'pointwise' equivalences.

We shall also speak of equivalences between 2-categories. By an *equivalence* between 2-categories $\mathfrak{K}$ and $\mathfrak{L}$, we mean a pair of functors $F: \mathfrak{K} \to \mathfrak{L}$ and $G: \mathfrak{L} \to \mathfrak{K}$, together with (specified) natural equivalences between the composites FG and GF and the respective identity functors. Occasionally it happens that we have not just equivalences but isomorphisms between these composites and the identities; if we wish to draw attention to the fact that this happens, we shall call (F, G) a *strong equivalence*.

When we speak of *representability* of a $\mathfrak{Cat}$-valued functor on a 2-category, we shall (in keeping with the general principle that diagrams in a 2-category cannot be expected to commute 'on the nose') interpret this notion in the 'up to natural equivalence' rather than the 'up to natural isomorphism' sense. Thus a representation of $F: \mathfrak{K} \to \mathfrak{Cat}$ consists of an object A of $\mathfrak{K}$ together with a natural equivalence between F and the (strict) functor $\mathfrak{K}(A, -)$ (here we are – temporarily – assuming that $\mathfrak{K}$ is locally small). Of course, the representing object A is unique up to equivalence in $\mathfrak{K}$, the equivalence between two representations A and A' being itself unique up to unique 2-isomorphism.

The idea of a representation can be formulated in elementary terms, even when $\mathfrak{K}$ is not locally small. For example, a *limit* for a diagram $D: \mathfrak{J} \to \mathfrak{K}$ consists of an object A of $\mathfrak{K}$ together with a (pseudo)cone over D with vertex A

(that is, a natural transformation from the constant diagram with value A to D) and an operation assigning to each cone over D with vertex B a morphism $B \to A$ and an invertible modification from the given cone to the composite of the limit cone with this morphism, and to each modification between cones with vertex B a transformation between the corresponding morphisms $B \to A$ such that an appropriate diagram commutes, ... We shall generally not spell out the details of such definitions; it is always possible to work them out for oneself, by asking what it would mean in elementary terms, in the locally small case, to have a natural equivalence between the appropriate $\mathfrak{Cat}$-valued functors, and then throwing away the requirement that $\mathfrak{K}$ be locally small.

It is by now well known that the ordinary 'conical' limits described in the last paragraph are not sufficient to capture the notion of completeness required for 2-categories. Instead, we need to consider weighted limits, which are representations of $\mathfrak{Cat}$-valued functors sending an object to the category of weighted cones over a given diagram (with a given weighting). (When they were first introduced by R. Street [1130], weighted limits were known as 'indexed limits'; but that term has been abandoned because of the clash with indexed categories, which we shall introduce in the next section.)

Definition 1.1.3

(*a*) A *weighted diagram* of type $\mathfrak{J}$ in a 2-category $\mathfrak{K}$ consists of a functor $D\colon \mathfrak{J} \to \mathfrak{K}$ together with a functor $W\colon \mathfrak{J} \to \mathfrak{Cat}$ (called the *weighting* of the diagram).

(*b*) A *weighted cone* over a weighted diagram (D, W) consists of an object A of $\mathfrak{K}$ (the *vertex* of the cone) together with

(i) for each object J of $\mathfrak{J}$, a functor $L(J)\colon W(J) \to \mathfrak{K}(A, D(J))$, and

(ii) for each morphism $f\colon J \to J'$ of $\mathfrak{J}$, a natural isomorphism $L(f)$ fitting into the diagram

$$\begin{tikzcd} W(J) \arrow[r, "L(J)"] \arrow[d, "W(f)"'] \arrow[dr, phantom, "\downarrow L(f)"] & \mathfrak{K}(A, D(J)) \arrow[d, "D(f)\circ(-)"] \\ W(J') \arrow[r, "L(J')"] & \mathfrak{K}(A, D(J')) \end{tikzcd}$$

such that the $L(f)$ are compatible in a suitable sense with composition in $\mathfrak{J}$, and such that for each 2-cell of $\mathfrak{J}$ an appropriate diagram of natural transformations commutes.

If L and L' are weighted cones over the same weighted diagram (D, W) with the same vertex A, then we may define a morphism $L \to L'$ to be a family of transformations $L(J) \to L'(J)$, $J \in \text{ob}\,\mathfrak{J}$, which are compatible with the $L(f)$ and $L'(f)$. Thus we have a category $\mathbf{Cone}(A, (D, W))$ of

cones over (D, W) with vertex A; and it is easy to see that the assignment $A \mapsto \mathbf{Cone}(A, (D, W))$ is contravariantly functorial in A.

(c) A *weighted limit* for the weighted diagram (D, W) is a representation of the functor just described; in other words, it consists of an object A equipped with a weighted cone L such that, for each object B of $\mathfrak{K}$, composition with L is part of an equivalence of categories $\mathfrak{K}(B, A) \simeq \mathbf{Cone}(B, (D, W))$. (We call a weighted limit *strong* if this equivalence is in fact an isomorphism for all B.)

Dually, given a functor $W: \mathfrak{J}^{\mathrm{op}} \to \mathfrak{Cat}$, we may define the notions of W-weighted cone *under* a diagram $D: \mathfrak{J} \to \mathfrak{K}$, and of a weighted colimit for the pair (D, W). (Thus weighted colimits in $\mathfrak{K}$ are the same thing as weighted limits in $\mathfrak{K}^{\mathrm{op}}$.)

Examples 1.1.4 (*a*) Ordinary diagrams in $\mathfrak{K}$ can be regarded as weighted diagrams for which W is the constant functor with value $\mathbf{1}$. Thus weighted limits include ordinary 'conical' limits as a special case. More generally, if the weighting functor $W: \mathfrak{J} \to \mathfrak{Cat}$ takes discrete categories as values, then a W-weighted diagram is equivalent to an ordinary diagram over a 'larger' 2-category, in which the Jth vertex of the original diagram is repeated as many times as there are objects of $W(J)$. (This is the reason why the difference between ordinary and weighted limits does not become apparent in the theory of 1-categories.)

(*b*) Suppose $\mathfrak{J}$ is the terminal category $\mathbf{1}$; then the weighting W simply picks out a particular (small) category $\mathcal{C}$, and a W-weighted limit for (the functor picking out) an object A of $\mathfrak{K}$ is what is commonly called a *cotensor* of A with $\mathcal{C}$; that is, an object $\mathcal{C} \pitchfork A$ together with a diagram of type $\mathcal{C}$ in $\mathfrak{K}(\mathcal{C} \pitchfork A, A)$, which is universal in the sense that composition with it induces an equivalence between $\mathfrak{K}(B, \mathcal{C} \pitchfork A)$ and the functor category $[\mathcal{C}, \mathfrak{K}(B, A)]$ for any B. (The corresponding weighted colimits are called *tensors* with $\mathcal{C}$, and denoted $\mathcal{C} \otimes A$.) If $\mathfrak{K}$ itself is $\mathfrak{Cat}$, then cotensors are just functor categories (and tensors are ordinary products), since we have $[\mathcal{C}, [\mathcal{B}, \mathcal{A}]] \cong [\mathcal{B} \times \mathcal{C}, \mathcal{A}] \cong [\mathcal{B}, [\mathcal{C}, \mathcal{A}]]$.

(*c*) Let $\mathfrak{J}$ be the finite category

$$\bullet \longrightarrow \bullet \longleftarrow \bullet$$

and let $W: \mathfrak{J} \to \mathfrak{Cat}$ be the diagram

$$\mathbf{1} \xrightarrow{\;0\;} \mathbf{2} \xleftarrow{\;1\;} \mathbf{1}$$

where, as usual, $\mathbf{n}$ denotes an n-element totally ordered set, and 0 and 1 denote distinct objects of $\mathbf{2}$. Then a W-weighted limit for a diagram

$$A \xrightarrow{\;f\;} C \xleftarrow{\;g\;} B$$

in $\mathfrak{K}$ is what is commonly called the *comma object* $(f \downarrow g)$ (also sometimes called a lax pullback – though this name properly belongs to a slightly different concept,

cf. 1.1.6 below); that is, it is a universal solution to the problem of completing the above diagram to one of the form

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] \arrow[dr, phantom, "\downarrow"] & A \arrow[d, "f"] \\ B \arrow[r, "g"] & C \end{tikzcd}$$

with a 2-cell as indicated. If $\mathfrak{K}$ has pullbacks and cotensors with $\mathbf{2}$, then it also has comma objects: $(f \downarrow g)$ may be constructed by forming the (iterated) pullback of the diagram

$$\begin{tikzcd} A \arrow[dr, "f"'] & & \mathbf{2} \pitchfork C \arrow[dl, "d_0"] \arrow[dr, "d_1"'] & & B \arrow[dl, "g"] \\ & C & & C & \end{tikzcd}$$

where d_0 and d_1 are the domain and codomain of the universal diagram of type $\mathbf{2}$ in $\mathfrak{K}(\mathbf{2} \pitchfork C, C)$. Conversely, if $\mathfrak{K}$ has comma objects then it has cotensors with $\mathbf{2}$, since we may take $\mathbf{2} \pitchfork C$ to be $(1_C \downarrow 1_C)$.

(*d*) Let $\mathfrak{J}$ be the finite category

$$\bullet \rightrightarrows \bullet$$

and give it the weighting represented by the diagram

$$\begin{tikzcd} \mathbf{1} \arrow[r, shift left, "0"] \arrow[r, shift right, "1"'] & \mathbf{2} \end{tikzcd}$$

in $\mathfrak{Cat}$. Then a weighted limit for a diagram

$$\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

in $\mathfrak{K}$ is what is commonly called an *inserter* for the pair (f, g); it is a universal solution to the problem of finding a morphism $h\colon C \to A$ together with a 2-cell $fh \to gh$. Inserters may be constructed from pullbacks, products and cotensors

with $\mathbf{2}$: specifically, if we form the pullback

$$\begin{tikzcd} I(f,g) \arrow[r] \arrow[d, "h"'] & \mathbf{2} \pitchfork B \arrow[d, "(d_0, d_1)"] \\ A \arrow[r, "(f,g)"] & B \times B \end{tikzcd}$$

then $I(f,g)$ is an inserter for the pair (f,g).

(*e*) To construct what are known as *inverters* as weighted limits, we need to take $\mathfrak{J}$ to be a 'real' 2-category (that is, one with non-identity 2-cells), namely

$$\begin{tikzcd} \bullet \arrow[r, bend left] \arrow[r, bend right] \arrow[r, phantom, "\downarrow"] & \bullet \end{tikzcd},$$

equipped with the weighting

$$\begin{tikzcd} \mathbf{1} \arrow[r, bend left, "0"] \arrow[r, bend right, "1"'] \arrow[r, phantom, "\downarrow"] & \mathbf{I} \end{tikzcd}$$

where $\mathbf{I}$ is the indiscrete category with two objects 0 and 1. An inverter for a diagram

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & B \end{tikzcd}$$

in $\mathfrak{K}$ is a universal solution to the problem of finding $h \colon C \to A$ such that $\alpha \circ h$ is invertible.

If we take the same $\mathfrak{J}$ as above, but replace W by the constant functor with value $\mathbf{1}$, we obtain the notion sometimes called an *identifier*; for a diagram

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & B \end{tikzcd}$$

in $\mathfrak{K}$, it is the universal solution to the problem of finding $h \colon C \to A$ such that $\alpha \circ h$ 'is' an identity 2-cell. At least, that is what we would get if we were working with strict limits; that is, if we replaced pseudo-cones by strict cones (ones in which the 2-cells $L(f)$ of Definition 1.1.3(*b*)(ii) are all identities). However, in the context of pseudo-cones which we have adopted as standard, the notion of identifier cannot be distinguished from that of inverter. More generally, *if W and $W' \colon \mathfrak{J} \to \mathfrak{Cat}$ are naturally equivalent functors, then W-weighted limits for diagrams $D \colon \mathfrak{J} \to \mathfrak{K}$ (whenever they exist) are also W'-weighted limits.*

Yet again, we can construct inverters from pullbacks and cotensors with $\mathbf{2}$: the inverter of α fits into the pullback square

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & B \arrow[d, "\ulcorner \iota_{1_B} \urcorner"] \\ A \arrow[r, "\ulcorner \alpha \urcorner"] & \mathbf{2} \pitchfork B \end{tikzcd}$$

where $\ulcorner \alpha \urcorner$ denotes the 1-cell into $\mathbf{2} \pitchfork B$ corresponding to the 2-cell α with codomain B.

(*f*) An *equifier* is the '2-dimensional analogue' of an equalizer: specifically, an equifier for a diagram

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\alpha\downarrow\ \downarrow\beta"] & B \end{tikzcd}$$

is a universal solution to the problem of finding $h\colon C \to A$ such that $\alpha \circ h = \beta \circ h$. We can describe equifiers as weighted limits: we take $\mathfrak{J}$ to be the 2-category indicated by the shape of the above diagram, equipped with the weighting

$$\begin{tikzcd} \mathbf{1} \arrow[r, bend left] \arrow[r, bend right] \arrow[r, phantom, "\downarrow\ \downarrow"] & \mathbf{2} \end{tikzcd}$$

where the two 2-cells are equal. And we can construct them from products, pullbacks and cotensors with $\mathbf{2}$: the equifier of α and β fits into the pullback square

$$\begin{tikzcd} \bullet \arrow[r] \arrow[d] & \mathbf{2} \pitchfork B \arrow[d, "\Delta"] \\ A \arrow[r, "(\ulcorner \alpha \urcorner, \ulcorner \beta \urcorner)"] & \mathbf{2} \pitchfork B \times \mathbf{2} \pitchfork B \end{tikzcd}$$

where Δ is the diagonal map.

Note also that inverters may be constructed from inserters and equifiers: to form the inverter of a 2-cell $\alpha\colon f \to g$, we may first 'insert' a universal 2-cell $\beta\colon gh \to fh$, and then successively form two equifiers to force the vertical composites $(\alpha \circ h)\beta$ and $\beta(\alpha \circ h)$ to equal the appropriate identity 2-cells. We remark in passing that the class of weighted limits which can be constructed from products, inserters and equifiers has received some attention [987].

We saw in Examples (*c*–*f*) above that many weighted limits can be constructed from cotensors with $\mathbf{2}$ and ordinary limits. This is typical: in the same

way that all finite ordinary limits can be constructed from finite products and equalizers (A1.2.1), we have

Theorem 1.1.5 *If a 2-category $\mathfrak{K}$ has finite (resp. small) products, equalizers and cotensors with* **2**, *then it has all finite (resp. small) weighted limits. (Here 'finite weighted limit' means one such that the index 2-category $\mathfrak{J}$ is finite, and the values of the weighting functor $W : \mathfrak{J} \to \mathfrak{Cat}$ are finite categories.)* □

We omit the proof, which may be found in [1130] or [583].

We shall also have occasion to consider lax and oplax limits: these too can be considered as a special case of weighted limits. By a *lax diagram* of shape $\mathfrak{J}$ in $\mathfrak{K}$, we of course mean a lax functor $\mathfrak{J} \to \mathfrak{K}$, as we defined the concept after 1.1.2. A *lax cone* over a lax diagram D is an oplax natural transformation from a constant diagram to D; that is, it assigns to each object J of $\mathfrak{J}$ a morphism $L(J): A \to D(J)$, and to each morphism $f: J \to J'$ a 2-cell $L(f)$ as in

$$\begin{tikzcd} A \arrow[r, "L(J)"] \arrow[dr, "L(J')"'] & D(J) \arrow[d, "D(f)"] \arrow[dl, phantom, "\downarrow L(f)"] \\ & D(J') \end{tikzcd}$$

subject to the usual conditions of coherence and compatibility with the 2-cells of $\mathfrak{J}$, but without the requirement that the $L(f)$ be invertible. An *oplax cone* is similar, except that the $L(f)$ point in the opposite direction. A *lax* (resp. *oplax*) *limit* for a lax diagram is of course a universal lax (resp. oplax) cone over it.

The notion of (op)lax limit is subsumed in that of a weighted limit, by the following result:

Lemma 1.1.6 *Any small 2-category $\mathfrak{J}$ has a 'relaxation' $\tilde{\mathfrak{J}}$ such that lax functors $D: \mathfrak{J} \to \mathfrak{K}$ correspond to ordinary functors $\tilde{D}: \tilde{\mathfrak{J}} \to \mathfrak{K}$. Moreover, $\tilde{\mathfrak{J}}$ has a canonical weighting W such that W-weighted cones over $\tilde{D}$ correspond to lax cones over D.* □

We omit the proof of 1.1.6, which is rather messy to write out in full generality, though the basic idea is straightforward enough. The underlying 1-category of $\tilde{\mathfrak{J}}$ is the free category generated by the underlying directed graph of the underlying 1-category of $\mathfrak{J}$ (that is, it has the same objects as $\mathfrak{J}$, and its morphisms are finite composable strings of morphisms of $\mathfrak{J}$), and its 2-cells are generated by the 2-cells of $\mathfrak{J}$ together with those whose images under a pseudofunctor defined

on $\tilde{\mathfrak{J}}$ will form the coherence 2-cells of the corresponding lax functor on $\mathfrak{J}$. The construction of the weighting W is similar.

Example 1.1.7 A particular case of lax diagrams and limits is of special interest: if $\mathfrak{J}$ is the terminal category $\mathbf{1}$, then a lax diagram of shape $\mathfrak{J}$ is simply a *monad* in $\mathfrak{K}$, i.e. an object A equipped with a morphism $t\colon A \to A$ and 2-cells $\eta\colon 1_A \to t$, $\mu\colon tt \to t$ satisfying the usual monad identities. A lax cone over such a diagram consists of a morphism $f\colon B \to A$ together with a 2-cell $\phi\colon tf \to f$ satisfying identities that look like those satisfied by an algebra for a monad; a universal lax cone (if it exists – in $\mathfrak{Cat}$ it may be taken to be the Eilenberg–Moore category of algebras for the monad, equipped with its forgetful functor to A) is called an *Eilenberg–Moore object* for the monad. Thus we see that Eilenberg–Moore objects are a particular case of weighted limits. (Note also that in this case the relaxation $\tilde{\mathbf{1}}$ may be identified with the strict monoidal category (= 2-category with one object) $\mathfrak{Ord}_f$, whose underlying 1-category is the additive monoid of natural numbers, and whose unique 'hom-category' is the category $\mathbf{Ord}_f$ of finite ordinals and order-preserving maps between them.)

In addition to adjunctions in a 2-category, we shall also wish to consider weak adjunctions: we say $f\colon A \to B$ is *weak left adjoint* to $g\colon B \to A$ if there exist 2-cells $\eta\colon 1_A \to gf$ and $\epsilon\colon fg \to 1_B$ such that

$$\begin{tikzcd} B \arrow[r, "g"] \arrow[dr, "1_B"'] & A \arrow[d, "f"] \arrow[dr, "1_A"] \arrow[dl, phantom, "\downarrow\epsilon"] \arrow[dr, phantom, "\downarrow\eta"'] & \\ & B \arrow[r, "g"'] & B \end{tikzcd}$$

pastes to the identity 2-cell 1_g (though the other triangular identity need not hold). Note that the assertions 'f is weak left adjoint to g' and 'g is weak right adjoint to f' are not equivalent: the latter would imply that the other triangular identity holds, though the one above need not.

Lemma 1.1.8 *Suppose $\mathfrak{K}$ is locally Cauchy-complete (i.e. idempotent 2-cells split in $\mathfrak{K}$). Then a morphism of $\mathfrak{K}$ has a weak left (right) adjoint iff it has a left (right) adjoint.*

Proof One direction is trivial. Conversely, suppose $g\colon B \to A$ has a weak left adjoint f as above. Then by considering the diagram

$$\begin{tikzcd}
A \arrow[d, "f"'] \arrow[dr, "1_A"] \arrow[dr, phantom, "\downarrow\eta"', shift right=3] & & \\
B \arrow[r, "g"] \arrow[dr, "1_B"'] & A \arrow[d, "f"] \arrow[dr, "1_A"] \arrow[dl, phantom, "\downarrow\epsilon", shift right=3] \arrow[dr, phantom, "\downarrow\eta"', shift right=3] & \\
& B \arrow[r, "g"] \arrow[dr, "1_B"'] & A \arrow[d, "f"] \arrow[dl, phantom, "\downarrow\epsilon", shift right=3] \\
& & B
\end{tikzcd}$$

we see that $(\epsilon \circ f)(f \circ \eta)$ is an idempotent 2-cell $f \to f$. Let

$$f \xrightarrow{\ \beta\ } h \xrightarrow{\ \alpha\ } f$$

be a splitting of this idempotent; then it is easy to verify that the 2-cells $\theta = (g \circ \beta)\eta\colon 1_A \to gh$ and $\phi = \epsilon(\alpha \circ g)\colon hg \to 1_B$ satisfy both triangular identities, and so make h left adjoint to g. □

Virtually all the 2-categories that we shall consider have the property that their idempotent 2-cells split. (It will be recalled that we verified this condition, for the particular 2-category $\mathfrak{Top}$, in A4.1.15.)

Weak adjunctions arise naturally in contexts where retracts are under consideration. Actually, we have to be slightly careful in defining the notions of 'idempotent' and 'retract' in a 2-category: we state the following definitions for future reference.

Definition 1.1.9 (*a*) By an *idempotent 1-cell* in a 2-category $\mathfrak{K}$, we mean a 1-cell $t\colon A \to A$ equipped with a 2-isomorphism $\mu\colon tt \to t$ such that $\mu \circ t$ and $t \circ \mu$ are equal as 2-cells $ttt \to tt$. (Equivalently, an idempotent may be regarded as a normalized pseudofunctor $M \to \mathfrak{K}$, where M is the two-element monoid $\{1, e\}$ with $e^2 = e$.)

(*b*) By a *splitting* of an idempotent (t, μ) in a 2-category $\mathfrak{K}$, we mean a pair of 1-cells $i\colon B \to A$, $r\colon A \to B$ together with 2-isomorphisms $\alpha\colon ri \to 1_B$ and $\beta\colon t \to ir$ such that μ equals the vertical composite $\beta^{-1}(i \circ \alpha \circ r)(\beta \circ \beta)$. We note that a splitting of an idempotent (t, μ) may be obtained as either a

limit or a colimit of the corresponding pseudo-functor $M \to \mathfrak{K}$; in particular, if $\mathfrak{K}$ has either equalizers or coequalizers (in the 2-categorical sense) then all its idempotents split. We say B is a *retract* of A if it occurs as a splitting of some idempotent endomorphism of A.

(c) By an *idempotent monad* in $\mathfrak{K}$, we of course mean a monad (A, t, η, μ) (as defined in 1.1.7) for which μ is an isomorphism. We note that (t, μ) is then necessarily an idempotent 1-cell as defined in (a) above, since the identity $\mu \circ t = t \circ \mu$ follows from the associative law plus the invertibility of μ. We note also that if (i, r, α, β) is a splitting of such an idempotent (t, μ), as defined in (b), then r is left adjoint to i: the unit of the adjunction is the composite $\beta\eta$, and the counit is the isomorphism α. (And in this case the splitting may alternatively be obtained as an Eilenberg–Moore object for the monad (A, t, η, μ).) We say an object B is an *adjoint retract* of A if it occurs as a splitting of an idempotent monad on A; dually, B is a *coadjoint retract* if it occurs as a splitting of an idempotent comonad.

We note that the equivalences (resp. the cores) in an allegory, as defined in A3.3.1, are exactly those symmetric idempotents which are idempotent monads (resp. comonads). (Of course, in a locally ordered 2-category, any monad is idempotent, since a 2-cell has a 2-sided inverse iff it has a 1-sided inverse.) We saw in A3.3.12 that, in order to split all the symmetric idempotents in an allegory, it is enough to split those which are either monads or comonads; in a general 2-category the corresponding result is of course false, but in Chapter C4 we shall encounter a number of important 2-categories where it suffices to split the idempotent comonads in order to split all idempotents. (Cf. also 1.1.15 below.)

We now return to the relationship between weak adjoints and retracts. The following result is typical:

Lemma 1.1.10 *Suppose given a commutative diagram*

$$\begin{tikzcd} A \arrow[r, "i"] \arrow[d, "f"] & B \arrow[r, "r"] \arrow[d, "g"] & A \arrow[d, "f"] \\ A' \arrow[r, "i'"] & B' \arrow[r, "r'"] & A' \end{tikzcd}$$

(up to isomorphism) in a 2-category $\mathfrak{K}$, such that $ri \cong 1_A$ and $r'i' \cong 1_{A'}$ by isomorphisms which are compatible in the obvious sense with the isomorphisms $fr \cong r'g$ and $gi \cong i'f$. Suppose further that g has a weak left adjoint. Then f has a weak left adjoint. In particular, if g has a left adjoint l and $\mathfrak{K}$ is locally Cauchy-complete, then f has a left adjoint m. Moreover, if the adjunction $(l \dashv g)$ is a coreflection (i.e. its unit is an isomorphism), then so is the adjunction $(m \dashv f)$.

Proof Let l denote the weak left adjoint of g, and consider the composite rli'. We have 2-cells

$$\alpha\colon 1_{A'} \cong r'i' \xrightarrow{r' \circ \eta \circ i'} r'gli' \cong frli'$$

and

$$\beta\colon rli'f \cong rlgi \xrightarrow{r \circ \epsilon \circ i} ri \cong 1_A,$$

where η and ϵ are the 'weak unit' and 'weak counit' of the weak adjunction between g and l. A straightforward diagram-chase shows that the composite $(f \circ \beta)(\alpha \circ f)$ is the identity 2-cell on f; so rli' is weak left adjoint to f. The second assertion follows from the first and 1.1.8. For the third, we note that if the unit $1 \to gl$ is an isomorphism, then so is the weak unit $1 \to frli'$; but the left adjoint m of f is obtained by splitting an idempotent endomorphism of rli', and the unit $1 \to fm$ is obtained by factoring $1 \to frli'$ through the split monomorphism $fm \rightarrowtail frli'$. It follows easily that both halves of this factorization must be isomorphisms. □

A particular context where weak adjoints occur is in the study of KZ-monads, which are a special class of (2-)monads on 2-categories playing a rôle somewhat analogous to that of idempotent monads in 1-category theory. By a monad $\mathbb{D} = (D, \eta, \mu)$ on a 2-category $\mathfrak{K}$, we of course mean a pseudo-monad; that is, D is a pseudofunctor $\mathfrak{K} \to \mathfrak{K}$, η and μ are pseudo-natural transformations, and the diagrams

$$\begin{tikzcd} D \arrow[r, "\eta_D"] \arrow[dr, "1"'] & DD \arrow[d, "\mu"] & D \arrow[l, "D\eta"'] \arrow[dl, "1"] \\ & D & \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} DDD \arrow[r, "D\mu"] \arrow[d, "\mu_D"'] & DD \arrow[d, "\mu"] \\ DD \arrow[r, "\mu"] & D \end{tikzcd}$$

commute not 'on the nose' but up to (specified, but nameless) invertible modifications, which are in turn required to satisfy appropriate coherence conditions. Similarly, when we speak of a $\mathbb{D}$-algebra, we mean a pseudo-algebra. (Monads on 2-categories are often called 'doctrines', but there seems no compelling reason to abandon the name familiar from 1-dimensional category theory.)

Definition 1.1.11 A monad $\mathbb{D} = (D, \eta, \mu)$ on a 2-category $\mathfrak{K}$ is called a *KZ-monad* if μ is naturally left adjoint to $\eta_D\colon D \to DD$, the counit of the adjunction being the invertible modification in the leftmost cell of the diagrams above.

KZ-monads were introduced independently by A. Kock [605] and V. Zöberlein [1262]; the name is apparently due to R. Street [1132]. Although the notion of KZ-monad has a 'handedness' (that is, it is not self-dual at the level of 2-cells), the apparent asymmetry in the definition, in that it refers to η_D but not to $D\eta$, is only apparent:

Lemma 1.1.12 *(D, η, μ) is a KZ-monad iff μ is naturally right adjoint to $D\eta$, the unit of the adjunction being the natural isomorphism appearing in the definition of a pseudo-monad.*

Proof Given the unit map $\phi_A : 1_{DDA} \to \eta_{DA}\mu_A$, we define the counit $\psi_A : (D\eta_A)\mu_A \to 1_{DDA}$ by pasting the diagram

$$\begin{tikzcd}
DDA \arrow[rrr, "\mu_A"] \arrow[dr, "DD\eta_A"] \arrow[dd, "1_{DDA}"'] & & & DA \arrow[ddd, "D\eta_A"] \\
& DDDA \arrow[dl, "D\mu_A"'] \arrow[dr, "1_{DDDA}"] \arrow[d, phantom, "\Downarrow D\phi_A"] & & \\
DDA \arrow[rr, "D\eta_{DA}"'] \arrow[drrr, bend right, "1_{DDA}"'] & {} & DDDA \arrow[dr, "\mu_{DA}"] & \\
& & & DDA
\end{tikzcd}$$

where the cells without explicit 2-arrows in them are those appearing in the definition of a pseudo-monad, plus an instance of the naturality of μ. It is clear from the construction that ψ_A is natural in A; the verification that it is indeed the counit of an adjunction ($\mu_A \dashv D\eta_A$) is tedious but straightforward. Conversely, given ψ_A, we obtain ϕ_A by pasting a similar diagram. □

KZ-monads may also be defined more symmetrically in terms of a modification $D\eta \to \eta_D$ satisfying certain axioms; see [605]. However, the importance of the notion resides in the existence of the adjunctions of 1.1.11 and 1.1.12.

Proposition 1.1.13 *Let $\mathbb{D}$ be a KZ-monad on a 2-category $\mathfrak{K}$, and A an object of $\mathfrak{K}$. Then*

(i) *Any retraction (up to 2-isomorphism) for $\eta_A : A \to DA$ is weak left adjoint to η_A.*

(ii) *Any $\mathbb{D}$-algebra structure on A is left adjoint to η_A.*

(iii) *Conversely, if η_A has a left adjoint such that the counit of the adjunction is an isomorphism, then this adjoint is a $\mathbb{D}$-algebra structure on A.*

Proof (i) Suppose given a morphism $f\colon DA \to A$ and a 2-isomorphism $\alpha\colon f\eta_A \to 1_A$. We obtain a 2-cell $\beta\colon 1_{DA} \to \eta_A f$ by pasting the diagram

$$\begin{tikzcd}
DA \arrow[r, "D\eta_A"] \arrow[dr, "1_{DA}"'] \arrow[rrrd, bend left=60, "1_{DA}"] \arrow[rrrd, phantom, bend left=30, "\Downarrow D\alpha^{-1}"] \arrow[ddr, "f"'] & DDA \arrow[d, "\mu_A"'] \arrow[dr, "1_{DDA}"] \arrow[d, phantom, shift left=6, "\Downarrow \phi_A"] & & \\
& DA \arrow[r, "\eta_{DA}"'] & DDA \arrow[r, "Df"] & DA \\
& A \arrow[rru, "\eta_A"'] & &
\end{tikzcd}$$

where ϕ is the unit of the adjunction between μ_A and η_{DA}, and the bottom cell commutes up to isomorphism by naturality of η. If we paste this to α, we obtain

$$\begin{tikzcd}
A \arrow[r, "\eta_A"] \arrow[dr, "\eta_A"'] \arrow[ddd, bend right, "1_A"'] & DA \arrow[dr, "D\eta_A"] \arrow[rrrdd, bend left=40, "1_{DA}"] \arrow[rrrdd, phantom, bend left=20, "\Downarrow D\alpha^{-1}"] & & & \\
& DA \arrow[r, "\eta_{DA}"] \arrow[dr, "1_{DA}"'] \arrow[dd, phantom, "\Downarrow \alpha"] & DDA \arrow[d, "\mu_A"'] \arrow[dr, "1_{DDA}"] \arrow[d, phantom, shift left=6, "\Downarrow \phi_A"] & & \\
& & DA \arrow[r, "\eta_{DA}"'] \arrow[d, "f"] & DDA \arrow[r, "Df"] & DA \\
& & A \arrow[rru, "\eta_A"'] & &
\end{tikzcd}$$

of which the two central cells collapse by one of the triangular identities for $(\mu_A \dashv \eta_{DA})$, and the rest then collapses to the identity 2-cell on η_A by naturality of η.

(ii) If we paste β to α the other way, we obtain

$$
\begin{tikzcd}
DA \arrow[r, "D\eta_A"] \arrow[dr, "1_{DA}"'] \arrow[rrrd, bend left=60, "1_{DA}"] \arrow[rrdd, bend right=60, "f"'] & DDA \arrow[d, "\mu"'] \arrow[dr, "1_{DDA}"] \arrow[dr, phantom, "\downarrow \phi_A"'] & & \downarrow D\alpha^{-1} \\
& DA \arrow[r, "\eta_{DA}"'] & DDA \arrow[r, "Df"] & DA \arrow[d, "f"] \\
& & A \arrow[ur, "\eta_A"] \arrow[r, "1_A"'] \arrow[r, phantom, bend left, "\alpha \downarrow"] & A
\end{tikzcd}
$$

which does not collapse to the identity in general; but if f is a $\mathbb{D}$-algebra structure, then we may modify the lower half of this diagram to yield

$$
\begin{tikzcd}
DA \arrow[r, "D\eta_A"] \arrow[dr, "1_{DA}"'] \arrow[rrrd, bend left=60, "1_{DA}"] & DDA \arrow[d, "\mu_A"'] \arrow[dr, "1_{DDA}"] \arrow[dr, phantom, "\downarrow \phi_A"'] & & \downarrow D\alpha^{-1} \\
& DA \arrow[r, "\eta_{DA}"] \arrow[dr, "1_{DA}"'] & DDA \arrow[r, "Df"] \arrow[d, "\mu_A"] & DA \arrow[d, "f"] \\
& & DA \arrow[r, "f"] & A
\end{tikzcd}
$$

of which two cells collapse by one of the triangular identities for $(\mu_A \dashv \eta_{DA})$, and the rest collapses to the identity 2-cell on f by the coherence conditions for a $\mathbb{D}$-algebra.

(iii) Conversely, suppose $\eta_A \colon A \to DA$ has a left adjoint f, such that the counit of the adjunction is an isomorphism $\alpha \colon f\eta_A \to 1_A$. Then $f \cdot Df$ is left

adjoint to $D\eta_A \cdot \eta_A$, whereas $f \cdot \mu_A$ is left adjoint to $\eta_{DA} \cdot \eta_A$; but the two right adjoints are isomorphic by the naturality of η, so we obtain an isomorphism between the two left adjoints. This latter, together with α, may be checked to satisfy the coherence conditions for a $\mathbb{D}$-algebra structure on A. □

Corollary 1.1.14 *If $\mathbb{D}$ is a KZ-monad on a 2-category $\mathfrak{K}$, then a given object A of $\mathfrak{K}$ admits (up to canonical isomorphism) at most one $\mathbb{D}$-algebra structure. If $\mathfrak{K}$ is locally Cauchy-complete, then such a structure exists iff the unit $\eta_A : A \to DA$ is split monic.* □

A further result on KZ-monads, which we shall exploit in Chapter C4, is the following:

Corollary 1.1.15 *Let $\mathbb{D}$ be a KZ-monad on a 2-category $\mathfrak{K}$, and (A, f) a $\mathbb{D}$-algebra in $\mathfrak{K}$. If f has a left adjoint in $\mathfrak{K}$, then (A, f) is expressible as a retract (in the 2-category $\mathfrak{K}^{\mathbb{D}}$ of $\mathbb{D}$-algebras) of a free algebra. The converse holds if $\mathfrak{K}$ is locally Cauchy-complete. Moreover, in the latter case, every retract of a free $\mathbb{D}$-algebra is a coadjoint retract of such an algebra.*

Proof If f has a left adjoint g, the latter is necessarily an algebra homomorphism $(A, f) \to (DA, \mu_A)$; for on taking right adjoints of all the 1-cells in the diagram

$$\begin{tikzcd} DA \arrow[r, "Dg"] \arrow[d, "f"'] & DDA \arrow[d, "\mu_A"] \\ A \arrow[r, "g"] & DA \end{tikzcd}$$

we obtain an instance of the naturality of η. But f is also a morphism of $\mathbb{D}$-algebras, by the 'associativity axiom' for an algebra; and the unit of the adjunction $(g \dashv f)$ is an isomorphism, since the counit of $(f \dashv \eta_A)$ is an isomorphism. So (A, f) is a retract of (DA, μ_A), and indeed a coadjoint retract since f is still right adjoint to g in $\mathfrak{K}^{\mathbb{D}}$.

Conversely, it follows from 1.1.12 that the algebra structure on a free $\mathbb{D}$-algebra has a left adjoint in $\mathfrak{K}$; and 1.1.8 ensures that this property is inherited by retracts. The last assertion follows from the other two. □

In Section B4.5 (and again in Section C4.3), we shall need to consider an 'intrinsic' notion of (co)completeness for objects of a 2-category. Once again, we take our motivation from the 2-category $\mathfrak{CAT}$. Recall that a category $\mathcal{C}$ is cocomplete iff left Kan extensions of $\mathcal{C}$-valued functors exist along any functor $G: \mathcal{D} \to \mathcal{E}$ between small categories, or more generally along any functor such that the comma categories $(G \downarrow A)$, $A \in \text{ob}\,\mathcal{E}$, are all small. Moreover, these Kan extensions are what are generally called 'pointwise Kan extensions': this means

that computation of the left Kan extension at an object A of $\mathcal{E}$ depends only on the comma category $(G \downarrow A)$. These ideas motivate the following definitions:

Definition 1.1.16 Let $\mathcal{A}$ be a class of morphisms in a 2-category $\mathfrak{K}$.

(a) We say an object D of $\mathfrak{K}$ is *$\mathcal{A}$-cocomplete* if, for every $f\colon A \to B$ in $\mathcal{A}$, the functor $f^{\#}\colon \mathfrak{K}(A, D) \to \mathfrak{K}(B, D)$ induced by composition with f has a left adjoint f_+.

(b) If $\mathfrak{K}$ has comma objects and, for any comma square

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\rightarrow"] & C \arrow[d, "g"] \\
A \arrow[r, "f"'] & B
\end{tikzcd}$$

in $\mathfrak{K}$, the condition $f \in \mathcal{A}$ implies $q \in \mathcal{A}$, then we say an object D is *pointwise $\mathcal{A}$-cocomplete* if it is $\mathcal{A}$-cocomplete and, for any comma square as above, the Beck–Chevalley condition holds, i.e. the canonical natural transformation $q_+p^{\#} \to g^{\#}f_+$ is an isomorphism.

(c) If $\mathfrak{K}$ has pullbacks and $\mathcal{A}$ is stable under pullback, then we say D is *linearly $\mathcal{A}$-cocomplete* if it satisfies a condition similar to (b), but with comma squares replaced by pullback squares.

(d) If $\mathcal{M}$ is a subclass of $\mathcal{A}$ (to be thought of as the class of monomorphisms in $\mathcal{A}$), then we say D is an *$\mathcal{A}$-cocomplete $\mathcal{M}$-injective* if it is $\mathcal{A}$-cocomplete and, for every $f\colon A \to B$ in $\mathcal{M}$, the unit map $1 \to f^{\#}f_+$ is an isomorphism.

(e) We say that a morphism $g\colon C \to D$ between $\mathcal{A}$-cocomplete objects of $\mathfrak{K}$ is *$\mathcal{A}$-cocontinuous* if, for any $f\colon A \to B$ in $\mathcal{A}$ and any $h\colon A \to C$, the canonical 2-cell $f_+(gh) \to g \cdot f_+(h)$ (where the f_+ on the left refers to the cocompleteness of D, and that on the right to C) is an isomorphism.

We note that this intrinsic notion of cocompleteness has at least one familiar property:

Lemma 1.1.17 *Suppose $g\colon C \to D$ has a right adjoint in $\mathfrak{K}$. Then it is $\mathcal{A}$-cocontinuous for any $\mathcal{A}$ such that C and D are $\mathcal{A}$-cocomplete.*

Proof $\mathcal{A}$-cocontinuity of g says that the square

$$\begin{tikzcd}
\mathfrak{K}(A, C) \arrow[r, "g_{\#}"] \arrow[d, "f_+"'] & \mathfrak{K}(A, D) \arrow[d, "f_+"] \\
\mathfrak{K}(B, C) \arrow[r, "g_{\#}"'] & \mathfrak{K}(B, D)
\end{tikzcd}$$

commutes up to isomorphism, where $g_\#$ is the functor 'compose with g'. But if g has a right adjoint k in $\mathfrak{K}$, then $(g_\# \dashv k_\#)$; and, on taking right adjoints of all the functors in the above square, we obtain one which commutes by associativity of composition in $\mathfrak{K}$. □

The connection between pointwise and linear $\mathcal{A}$-cocompleteness is given by the following result.

Lemma 1.1.18 *Suppose $\mathfrak{K}$ has pullbacks and cotensors with $\mathbf{2}$, and let $\mathcal{A}$ be a class of morphisms of $\mathfrak{K}$ which is 'stable under transfer across comma squares' in the sense of* 1.1.16(*b*). *Let $\mathcal{B}$ be the class of all morphisms q of $\mathfrak{K}$ which occur as the 'upper edge' of a comma square*

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\rightarrow"] & C \arrow[d, "g"] \\
A \arrow[r, "f"] & B
\end{tikzcd}$$

whose 'lower edge' f is in $\mathcal{A}$. Then $\mathcal{B}$ is stable under pullback, and an object D of $\mathfrak{K}$ is pointwise $\mathcal{A}$-cocomplete iff it is linearly $\mathcal{B}$-cocomplete.

Proof The first assertion is immediate from the construction of comma objects from cotensors and pullbacks (cf. 1.1.4(*c*)): if we have a diagram

$$\begin{tikzcd}
P \arrow[r] \arrow[d] & D \arrow[d] \\
(f \downarrow g) \arrow[r] \arrow[d] \arrow[dr, phantom, "\rightarrow"] & C \arrow[d] \\
A \arrow[r] & B
\end{tikzcd}$$

where the upper square is a pullback and the lower one is a comma square, then the whole is a comma square.

Suppose now that D is linearly $\mathcal{B}$-cocomplete; let $f: A \to B$ be a morphism in $\mathcal{A}$, and form the comma square

$$\begin{tikzcd}
(f \downarrow 1_B) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\to"] & B \arrow[d, "1_B"] \\
A \arrow[r, "f"] & B
\end{tikzcd}$$

We claim first that p has a left adjoint in $\mathfrak{K}$: the morphisms $1_A: A \to A$ and $f: A \to B$, together with the identity 2-cell $f \to f$, induce a morphism $g: A \to (f \downarrow 1_B)$ such that $pg \cong 1_A$ and $qg \cong f$, and the universal 2-cell $qgp \cong fp \to q$, together with the isomorphism $pgp \cong p$, induce a 2-cell $gp \to 1$ which is the counit of an adjunction $(g \dashv p)$ (the unit being the isomorphism $1 \cong pg$). Thus the functor $p^{\#}: \mathfrak{K}(A, D) \to \mathfrak{K}((f \downarrow 1_B), D)$ is left adjoint to $g^{\#}$, and so we can define the left adjoint of $f^{\#} \cong g^{\#} q^{\#}$ to be the composite $q_+ p^{\#}$. Thus D is $\mathcal{A}$-cocomplete; to see that it is pointwise $\mathcal{A}$-cocomplete, we simply factor a comma square whose bottom edge is in $\mathcal{A}$ as a pullback square whose top and bottom edges are in $\mathcal{B}$, followed by a comma square as above whose right-hand edge is an identity morphism.

Conversely, suppose D is pointwise $\mathcal{A}$-cocomplete. Then it is certainly $\mathcal{B}$-cocomplete, since $\mathcal{B}$ is a subclass of $\mathcal{A}$; the only problem is to show linearity. Suppose given a pullback square

$$\begin{tikzcd}
P \arrow[r, "k"] \arrow[d, "h"'] & C \arrow[d, "g"] \\
A \arrow[r, "f"] & B
\end{tikzcd}$$

with $f \in \mathcal{B}$; then by definition we have a comma square

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "a"'] \arrow[dr, phantom, "\xrightarrow{\alpha}"] & B \arrow[d, "b"] \\
A' \arrow[r, "f'"] & B'
\end{tikzcd}$$

with $f' \in \mathcal{A}'$, so we may regard P as the comma object $(f' \downarrow bg)$. We may also construct the comma object

$$\begin{tikzcd} (f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\xrightarrow{\beta}"] & C \arrow[d, "g"] \\ A \arrow[r, "f"'] & B \end{tikzcd} \quad ;$$

now we have a comparison map $u: P \to (f \downarrow g)$ defined by setting $pu \cong h$, $qu \cong k$ and taking the 2-cell $fpu \to gqu$ to be the canonical isomorphism $fh \cong gk$. But we also have a comparison map $v: (f \downarrow g) \to (f' \downarrow bg) \cong P$ defined by setting $ahv \cong ap$, $kv \cong q$ and by taking the 2-cell $\gamma: f'ahv \cong f'ap \to bgq \cong bgkv$ to be that obtained by pasting the 2-cells α and β in the two comma squares above. Moreover, we have $kvu \cong qu \cong k$, $ahvu \cong apu \cong ah$ and these isomorphisms identify $\gamma \circ u$ with $\alpha \circ h$, so $vu \cong 1_P$; and $\beta: fp \to gq \cong gkv \cong fhv \cong fpuv$ and the isomorphism $ap \cong ahv \cong apuv$ induce a 2-cell $p \to puv$, which in turn can be combined with the isomorphism $q \cong kv \cong quv$ to yield a 2-cell $1_{(f \downarrow g)} \to uv$. It is again easy to see that this 2-cell, and the isomorphism $vu \cong 1_P$, define an adjunction $(v \dashv u)$, so that $(u^{\#} \dashv v^{\#})$. Now we have $k_+ h^{\#} \cong k_+ u^{\#} p^{\#} \cong q_+ p^{\#}$, since $k_+ u^{\#}$ is left adjoint to $v^{\#} k^{\#} \cong q^{\#}$; so the Beck–Chevalley condition for the pullback square follows from that for the comma square. □

Concerning the relation between cocompleteness and injectivity, we note the following. We say a morphism $f: A \to B$ in a 2-category $\mathfrak{K}$ is *fully monic* if, for any object C of $\mathfrak{K}$, composition with f induces a full and faithful functor $\mathfrak{K}(C, A) \to \mathfrak{K}(C, B)$. For example, inclusions are fully monic in $\mathfrak{Top}$, since if f_* is full and faithful then any natural transformation $f_* g_* \to f_* h_*$ must be of the form $f_* \circ \alpha$ for a unique $\alpha: g_* \to h_*$. (Similarly, geometric morphisms with full and faithful inverse image functors – which we shall christen *connected morphisms* in C1.5.7 – are fully epic in $\mathfrak{Top}$.)

Lemma 1.1.19 *Suppose $\mathcal{A}$ is a class of morphisms of $\mathfrak{K}$ such that D is either pointwise or linearly $\mathcal{A}$-cocomplete. Then D is $\mathcal{M}$-injective, where $\mathcal{M}$ is the class of full monomorphisms in $\mathcal{A}$.*

Proof If $f: A \to B$ is fully monic, then it is easy to see that the square

$$\begin{tikzcd} A \arrow[r, "1_A"] \arrow[d, "1_A"'] & A \arrow[d, "f"] \\ A \arrow[r, "f"'] & B \end{tikzcd}$$

is a pullback in $\mathfrak{K}$; it is not a comma square in general, but the comma object $(f \downarrow f)$ coincides with $(1_A \downarrow 1_A)$. So the isomorphism $f^\# f_+ \cong 1$ follows in either case from the isomorphisms $1_{A+} \cong 1_A^\# \cong 1$. □

Suggestions for further reading: Bénabou [99], Bird et al. [120], Bunge & Funk [199], Kelly [583], Kelly et al. [585], Kock [605], Power & Robinson [987], Street [1130, 1132].

B1.2 Indexed categories

In the non-elementary parts of the theory of categories, we frequently find it necessary to talk about set-indexed families of objects or morphisms of a given category. Bearing in mind the viewpoint that a topos is a category of sets, the study of category theory relative to a topos $\mathcal{S}$ is going to demand that we should be able to talk about families of objects indexed by an object of $\mathcal{S}$. The theory of indexed categories, and the (essentially equivalent) theory of fibrations which we shall expound in the next section, were created to provide a way of doing this.

The important structure which indexed families possess is the possibility of re-indexing along a map: if $x: I \to J$ is a morphism in **Set** and $(A_j \mid j \in J)$ is a J-indexed family of objects of a category $\mathcal{C}$, we should be able to form the I-indexed family $(A_{x(i)} \mid i \in I)$. Thus the definition of an $\mathcal{S}$-indexed category requires not only the specification of a category $\mathcal{C}^I$ of I-indexed families for each object I of $\mathcal{S}$, but also of a re-indexing functor $x^*: \mathcal{C}^J \to \mathcal{C}^I$ for each $x: I \to J$ in $\mathcal{S}$. (For the rest of Part B, we shall as far as possible reserve the letters I, J, K and x, y, z respectively for objects and morphisms of the base category $\mathcal{S}$.) Our experience of the case $\mathcal{S} = \mathbf{Set}$ might lead us to expect that the assignment $x \mapsto x^*$ should be (contravariantly) functorial on $\mathcal{S}$; but the principles enunciated in the last section encourage us to make it merely pseudofunctorial, and this is supported by the examples we shall have occasion to consider.

In order to effect a conceptual simplification in the definitions which follow, we shall use our informal notation $\mathfrak{CAT}$ for the meta-2-category of categories, as explained in the last section; and we shall identify our base category $\mathcal{S}$ with the corresponding locally discrete 2-category. For the purposes of the definition, $\mathcal{S}$ is not required to have any structure beyond that of a category, but in all the examples of interest to us it will be at least cartesian.

Definition 1.2.1 (*a*) An *$\mathcal{S}$-indexed category* $\mathbb{C}$ is a pseudofunctor $\mathcal{S}^{\mathrm{op}} \to \mathfrak{CAT}$. By convention, we write $\mathcal{C}^I$ for the category $\mathbb{C}(I)$ ($I \in \mathrm{ob}\ \mathcal{S}$) and $x^*: \mathcal{C}^J \to \mathcal{C}^I$ for the functor $\mathbb{C}(x)$ ($x: I \to J$ in $\mathcal{S}$). If necessary, we shall write $\theta_{x,y}$ for the coherence isomorphism $x^* y^* \to (yx)^*$ (where (x, y) is a composable pair of morphisms of $\mathcal{S}$), but in general we shall suppress all mention of these isomorphisms whenever it is safe to do so. If $\mathcal{S}$ has a terminal object 1, we shall call $\mathcal{C}^1$ the *underlying ordinary category* of the $\mathcal{S}$-indexed category $\mathbb{C}$.

(*b*) If $\mathbb{C}$ and $\mathbb{D}$ are $\mathcal{S}$-indexed categories, an *$\mathcal{S}$-indexed functor* $F\colon \mathbb{C} \to \mathbb{D}$ is a pseudo-natural transformation; that is, it assigns to each object I of $\mathcal{S}$ a functor $F^I : \mathcal{C}^I \to \mathcal{D}^I$ and to each morphism $x\colon I \to J$ a natural isomorphism ϕ_x as in

$$\begin{tikzcd} \mathcal{C}^J \arrow[r, "x^*"] \arrow[d, "F^J"'] \arrow[dr, phantom, "\downarrow \phi_x"] & \mathcal{C}^I \arrow[d, "F^I"] \\ \mathcal{D}^J \arrow[r, "x^*"] & \mathcal{D}^I \end{tikzcd}$$

(again, we shall suppress the name of ϕ_x whenever possible), subject to the appropriate compatibility conditions with the $\theta_{x,y}$.

(*c*) If F and $G\colon \mathbb{C} \to \mathbb{D}$ are both $\mathcal{S}$-indexed functors, an *$\mathcal{S}$-indexed natural transformation* $\alpha\colon F \to G$ is a modification; that is, it assigns to each $I \in \text{ob}\, \mathcal{S}$ a natural transformation $\alpha^I : F^I \to G^I$, subject to compatibility conditions with the coherence isomorphisms ϕ_x and γ_x of F and G.

For a given $\mathcal{S}$, the $\mathcal{S}$-indexed categories, functors and natural transformations clearly form a (meta-) 2-category, which we shall denote by $\mathfrak{CAT}_{\mathcal{S}}$.

Examples 1.2.2 (*a*) Any ordinary category $\mathcal{C}$ may be made the underlying ordinary category of a **Set**-indexed category $\mathbb{C}$, by setting $\mathcal{C}^I$ to be the product of I copies of $\mathcal{C}$ (i.e. the category of I-indexed families of objects of $\mathcal{C}$, as classically understood). We shall refer to this as the *naive indexing* of $\mathcal{C}$ over **Set**. Similarly, ordinary functors and natural transformations extend to **Set**-indexed functors and natural transformations between the naive indexings of the corresponding categories.

(*b*) However, even over **Set** an indexed category is not determined by its underlying ordinary category; there may be 'non-naive' ways of indexing the latter over **Set**. For example, if $\mathcal{C} = \mathbf{Set}_f$ we could define $\mathcal{C}^I$ to consist of those I-indexed families $(A_i \mid i \in I)$ of finite sets such that the cardinality of A_i is bounded (independently of i), and all I-indexed families of functions between them. More radically, for any $\mathcal{C}$ we could define $\mathcal{C}^I$ to consist only of constant I-indexed families of objects of $\mathcal{C}$. In Section B1.5 we shall investigate 'stack conditions' which may be used to distinguish the naive indexing from other possible indexings of a given category.

(*c*) An example which will be of fundamental importance in all that follows: let $\mathcal{S}$ be a cartesian category, and define an $\mathcal{S}$-indexed category $\mathbb{S}$ by $\mathcal{S}^I = \mathcal{S}/I$, with x^* taken to be the functor 'pullback along x' already denoted thus in A1.2.8. (Recall our convention that a category with pullbacks comes equipped with a choice of 'canonical' pullbacks (cf. A1.2.1), so we do indeed have such a functor x^* for each morphism x of $\mathcal{S}$. However, we do not expect the composite of two canonical pullback squares to be canonical, so the assignment $x \mapsto x^*$ will not be strictly functorial – as it was in examples (*a*) and (*b*).) We note that

the underlying ordinary category of $\mathbb{S}$ is (isomorphic to) $\mathcal{S}$, and we call $\mathbb{S}$ the *canonical indexing* of $\mathcal{S}$ over itself. In the case $\mathcal{S} = \mathbf{Set}$, A1.1.6 assures us that the canonical indexing is equivalent (in $\mathfrak{CAT}_{\mathbf{Set}}$) to the naive indexing.

(*d*) More generally than (*c*), if we are given a functor $F\colon \mathcal{S} \to \mathcal{C}$ and $\mathcal{C}$ is cartesian, then we may form an indexed category $\mathbb{C}$ over $\mathcal{S}$ by setting $\mathcal{C}^I = \mathcal{C}/F(I)$ and $x^* =$ pullback along $F(x)$. If $\mathcal{S}$ has and F preserves 1, then the underlying ordinary category of $\mathbb{C}$ is isomorphic to $\mathcal{C}$; and if F preserves pullbacks, then it extends to an $\mathcal{S}$-indexed functor $\mathbb{S} \to \mathbb{C}$ in an obvious way. In fact, if $\mathcal{S}$ is cartesian and $\mathfrak{CART}$ denotes the meta-2-category of cartesian categories, cartesian functors and natural transformations, then this construction defines a 2-functor from the co-slice category $\mathcal{S}\backslash\mathfrak{CART}$ to $\mathfrak{CAT}_{\mathcal{S}}$.

(*e*) Example (*d*) is itself a special case of a more general construction, namely change of base. If $F\colon \mathcal{T} \to \mathcal{S}$ is a functor and $\mathbb{C}$ is an $\mathcal{S}$-indexed category, we obtain a $\mathcal{T}$-indexed category $F^*\mathbb{C}$ by setting $(F^*\mathcal{C})^I = \mathcal{C}^{F(I)}$ and $x^* = (F(x))^*$. This construction also works on $\mathcal{S}$-indexed functors and natural transformations; thus the assignment $\mathcal{S} \mapsto \mathfrak{CAT}_{\mathcal{S}}$ itself becomes (contravariantly) functorial on $\mathfrak{CAT}$. (It is even 2-functorial, although it is also contravariant at the level of 2-cells: a natural transformation $\alpha\colon F \to G$ yields an indexed natural transformation $G^*\mathbb{C} \to F^*\mathbb{C}$ whose I-component is $(\alpha_I)^*\colon \mathcal{C}^{G(I)} \to \mathcal{C}^{F(I)}$.)

(*f*) A special case of (*e*) which will be of particular importance: if $\mathcal{S}$ has finite products and F is the functor $(-) \times I\colon \mathcal{S} \to \mathcal{S}$, then we shall write $\mathbb{C}^I$ for $F^*\mathbb{C}$, where $\mathbb{C}$ is an $\mathcal{S}$-indexed category. That is, $(\mathcal{C}^I)^J = \mathcal{C}^{J\times I}$, and if $x\colon J \to K$ then $x^*\colon (\mathcal{C}^I)^K \to (\mathcal{C}^I)^J$ is $(x \times 1_I)^*\colon \mathcal{C}^{K\times I} \to \mathcal{C}^{J\times I}$. The assignment $I \mapsto \mathbb{C}^I$ is itself contravariantly functorial in I (as well as covariantly 2-functorial in $\mathbb{C}$). In particular, we have an $\mathcal{S}$-indexed functor $\Delta\colon \mathbb{C} \to \mathbb{C}^I$ defined by $\Delta^J = \pi_1^*\colon \mathcal{C}^J \to \mathcal{C}^{J\times I}$.

(*g*) Once again, suppose $\mathcal{S}$ is cartesian. Since monomorphisms in $\mathcal{S}$ are stable under pullback, the assignment $I \mapsto \mathrm{Sub}_{\mathcal{S}}(I)$ defines an indexed category, in fact an indexed subcategory of the category $\mathbb{S}$ of example (*c*), which we shall denote by $\mathbb{S}\mathrm{ub}$ (or $\mathbb{S}\mathrm{ub}_{\mathcal{S}}$, if it is necessary to specify $\mathcal{S}$). As we shall see, many of the categorical properties which we studied in Chapter A1 can be interpreted as properties of this indexed category.

(*h*) The following example will not be of major importance for us (except as a source of counterexamples), but it lies at the heart of the extensive subject known as categorical topology. Given a set I, let $T(I)$ denote the set of all topologies on I, regarded as a preorder (in fact a complete lattice) with the *opposite* of the inclusion ordering – that is, there is a morphism $\mathcal{U} \to \mathcal{U}'$ in $T(I)$ iff $\mathcal{U}$ is finer than $\mathcal{U}'$. If $x\colon I \to J$ is a function and $\mathcal{V} \in T(J)$, there is a unique coarsest topology $x^*\mathcal{V}$ on I for which x is continuous, namely

$$x^*\mathcal{V} = \{x^{-1}(V) \mid V \in \mathcal{V}\};$$

this defines an order-preserving map $x^*\colon T(J) \to T(I)$, and so makes the assignment $I \mapsto T(I)$ into a Set-indexed category $\mathbb{T}$. Note that, for each $x\colon I \to J$, x^*

has a left adjoint x_*, which sends a topology $\mathcal{U}$ on I to the finest topology on J making x continuous, i.e.

$$x_*\mathcal{U} = \{V \subseteq J \mid x^{-1}(V) \in \mathcal{U}\}.$$

We shall return to the discussion of such adjoints in Section B1.4 below. Other types of 'topological' structures that can be imposed on sets – uniformities, proximities and the like – may also be collected into **Set**-indexed preorders of this kind; it is a slight exaggeration, but not a vast one, to say that categorical topology is precisely the study of these **Set**-indexed preorders.

A question which will frequently be of interest to us is whether a given ordinary functor between the underlying ordinary categories of two indexed categories can be extended to an indexed functor. In this connection the following result, though seemingly rather special, is often useful:

Lemma 1.2.3 *Let $F: \mathcal{S} \to \mathcal{C}$ be a functor between cartesian categories, having a right adjoint R. Then R has a canonical extension to an $\mathcal{S}$-indexed functor $\mathbb{C} \to \mathbb{S}$, where $\mathbb{S}$ is the canonical indexing of $\mathcal{S}$ over itself, as defined in* 1.2.2(*c*), *and $\mathbb{C}$ is defined as in* 1.2.2(*d*).

Proof For each object I of $\mathcal{S}$, let R^I denote the composite

$$\mathcal{C}/FI \xrightarrow{R/FI} \mathcal{S}/RFI \xrightarrow{\eta_I^*} \mathcal{S}/I$$

where η is the unit of $(F \dashv R)$. To verify that this is an indexed functor, let $x: I \to J$ be a morphism of $\mathcal{S}$, and consider the diagram

$$\begin{tikzcd}
\mathcal{C}/FJ \arrow[r, "R/FJ"] \arrow[d, "(Fx)^*"] & \mathcal{S}/RFJ \arrow[r, "\eta_J^*"] \arrow[d, "(RFx)^*"] & \mathcal{S}/J \arrow[d, "x^*"] \\
\mathcal{C}/FI \arrow[r, "R/FI"] & \mathcal{S}/RFI \arrow[r, "\eta_I^*"] & \mathcal{S}/I
\end{tikzcd}$$

Here the left-hand square commutes up to isomorphism because R, being a right adjoint, preserves pullbacks, and the right-hand one commutes by naturality of η. □

In the case where F preserves pullbacks, and so extends (as we saw in 1.2.2(*d*)) to an $\mathcal{S}$-indexed functor $\mathbb{S} \to \mathbb{C}$, it is easy to verify that R is an indexed right adjoint to F – that is, for each object I we have an adjunction $(F^I \dashv R^I)$, and the units and counits of these adjunctions form $\mathcal{S}$-indexed natural transformations.

Remark 1.2.4 It is also easy to verify that if F has a left adjoint L, then F^I has a left adjoint L^I for each I; explicitly, L^I sends an object $(A \to FI)$ of

$\mathcal{C}/FI$ to its transpose $(LA \to I)$ in $\mathcal{S}/I$. However, the L^I do not in general form an $\mathcal{S}$-indexed functor $\mathbb{C} \to \mathbb{S}$; note in particular that the commutativity up to isomorphism of the square

$$\begin{tikzcd} \mathcal{C} \arrow[r, "(FI)^*"] \arrow[d, "L"] & \mathcal{C}/FI \arrow[d, "L^I"] \\ \mathcal{S} \arrow[r, "I^*"] & \mathcal{S}/I \end{tikzcd} ,$$

which would be required for the L^I to form an indexed functor, is just the 'Frobenius reciprocity' condition $L(A \times FI) \cong LA \times I$ of A1.5.8. We may thus deduce that *if $\mathcal{S}$ and $\mathcal{C}$ are locally cartesian closed categories and $F: \mathcal{S} \to \mathcal{C}$ is a functor having a left adjoint L, then the $\mathcal{S}$-indexed functor $\mathbb{S} \to \mathbb{C}$ of* 1.2.2(*d*) *has an indexed left adjoint iff F is locally cartesian closed (i.e. $F/I: \mathcal{S}/I \to \mathcal{C}/FI$ is a cartesian closed functor for each I).*

When dealing with $\mathcal{S}$-indexed categories, we naturally wish to be able to handle not just individual indexed functors and natural transformations between them, but also I-indexed families of such things, where I is any object of $\mathcal{S}$. Provided $\mathcal{S}$ has finite products, the construction of 1.2.2(f) equips us with the means of doing so: we define an I-indexed family of functors $\mathbb{C} \to \mathbb{D}$ to be a single indexed functor $\mathbb{C} \to \mathbb{D}^I$, and similarly for natural transformations. If $x: I \to J$ in $\mathcal{S}$, we have an indexed functor $x^*: \mathbb{D}^J \to \mathbb{D}^I$, whose K-component is $(1_K \times x)^*: \mathcal{D}^{K\times J} \to \mathcal{D}^{K \times I}$; these functors are themselves pseudo-functorial in x, and so composition with them can be used to define 're-indexing' for families of functors: that is, if we write $[\mathbb{C}, \mathbb{D}]$ for the ordinary category of indexed functors $\mathbb{C} \to \mathbb{D}$ and indexed natural transformations between them, then the functors $(F \mapsto x^* \circ F): [\mathbb{C}, \mathbb{D}^J] \to [\mathbb{C}, \mathbb{D}^I]$ form the transition maps of an $\mathcal{S}$-indexed category which we denote $[\![\mathbb{C}, \mathbb{D}]\!]$, the *indexed functor category* of indexed functors $\mathbb{C} \to \mathbb{D}$.

We shall be mainly concerned with indexed functor categories $[\![\mathbb{C}, \mathbb{D}]\!]$ when $\mathbb{C}$ is an internal category, a concept that we shall meet in Section B2.3 below. However, the following simple example is instructive:

Example 1.2.5 Let I be an object of $\mathcal{S}$, and let $\mathbb{I}$ denote the $\mathcal{S}$-indexed category such that $\mathcal{I}^J$ is the discrete category whose objects are all morphisms $J \to I$ in $\mathcal{S}$, and such that if $x: K \to J$ in $\mathcal{S}$ then $x^*: \mathcal{I}^J \to \mathcal{I}^K$ is the operation of composition with x. Given any $\mathcal{S}$-indexed category $\mathbb{D}$, each object A of $\mathcal{D}^I$ defines an indexed functor $\mathbb{I} \to \mathbb{D}$ sending $y: J \to I$ to $y^*A \in \mathrm{ob}\, \mathcal{D}^J$; but every indexed functor $F: \mathbb{I} \to \mathbb{D}$ is (canonically) isomorphic to the one obtained in this way from the object $F^I(1_I)$ of $\mathcal{D}^I$. Thus we have an equivalence between $\mathcal{D}^I$ and the functor category $[\mathbb{I}, \mathbb{D}]$; more generally, $[\mathbb{I}, \mathbb{D}^J]$ is equivalent to $\mathcal{D}^{I\times J}$ (and

hence to $\mathcal{D}^{J\times I}$). So we may conclude that the indexed functor category $[\![\mathbb{I},\mathbb{D}]\!]$ is equivalent to $\mathbb{D}^I$, as defined in 1.2.2(f).

We may also define indexed families of functors $\mathcal{C} \to \mathbb{D}$ when $\mathcal{C}$ is an ordinary category and $\mathbb{D}$ is an $\mathcal{S}$-indexed category: in this case, an I-indexed family of functors means simply an ordinary functor $\mathcal{C} \to \mathcal{D}^I$. Once again, these may be made into (the I-indexed families of objects of) an $\mathcal{S}$-indexed category, which we denote by $[\![\mathcal{C},\mathbb{D}]\!]$. It is straightforward to verify that if $\mathbb{C}$ and $\mathbb{D}$ are the naive **Set**-indexings of ordinary categories $\mathcal{C}$ and $\mathcal{D}$, then both $[\![\mathcal{C},\mathbb{D}]\!]$ and $[\![\mathbb{C},\mathbb{D}]\!]$ are equivalent to the naive indexing of the ordinary functor category $[\mathcal{C},\mathcal{D}]$.

Suggestion for further reading: Paré & Schumacher [933].

B1.3 Fibred categories

The notion of indexed category, though conceptually clear and attractive, is not exactly an elementary one: it is a collection of (potentially) 'large' entities indexed by a 'large' entity, and such things are hard to handle in set-theoretic metatheories, unless we make powerful assumptions such as the existence of 'Grothendieck universes'. However, there is a way of combining all the data in an indexed category into a single entity (itself due to Grothendieck, and commonly called the *Grothendieck construction*), which leads to an elementary theory which is equivalent (in all essential aspects) to that of indexed categories. We devote the present section to studying some aspects of this theory.

Definition 1.3.1 Let $\mathcal{S}$ be a category, and $\mathbb{C}$ an indexed category over $\mathcal{S}$. We define the *Grothendieck category* $\mathcal{G}(\mathbb{C})$ of $\mathbb{C}$ as follows: its objects are pairs (I, A) where I is an object of $\mathcal{S}$ and A is an object of $\mathcal{C}^I$, and morphisms $(I, A) \to (J, B)$ are pairs (x, f) where $x\colon I \to J$ in $\mathcal{S}$ and $g\colon A \to x^*(B)$ in $\mathcal{C}^I$. The composite of (x, f) and $(y, g)\colon (J, B) \to (K, C)$ is (yx, h), where h is the composite

$$A \xrightarrow{\ f\ } x^*(B) \xrightarrow{x^*(g)} x^*y^*(C) \xrightarrow{\theta_{x,y}(C)} (yx)^*(C);$$

the coherence conditions for the $\theta_{x,y}$ ensure that this composition is associative, and so $\mathcal{G}(\mathbb{C})$ is a category. There is an obvious functor $\Pi\colon \mathcal{G}(\mathbb{C}) \to \mathcal{S}$ sending (I, A) to I and (x, f) to x.

We should remark that a particular case of the Grothendieck construction (albeit for a covariant rather than a contravariant functor) has been seen much earlier on in this work: in fact in the proof of A1.1.7. If we are given a functor $F\colon \mathcal{S} \to$ **Set**, and we regard it as a $\mathfrak{Cat}$-valued (pseudo)functor taking discrete categories as values, then the Grothendieck construction applied to it yields the category $\mathcal{F}$ which we considered there. We shall have more to say about this special case later on.

The fibre of Π over the object I (i.e. the non-full subcategory of $\mathcal{G}(\mathbb{C})$ consisting of the objects mapping to I and the morphisms mapping to 1_I) is just

(an isomorphic copy of) the category C^I, and we shall identify it therewith. It is convenient to picture the base category S as spread out horizontally and the fibres C^I as running vertically above it, as in the following picture:

$$
\begin{tikzcd}
C^I & C^J & \\
A \arrow[d, "f"'] \arrow[dr, "{(x,f)}"] & & \mathcal{G}(\mathbb{C}) \\
x^*(B) \arrow[r, "{(x,1)}"'] & B & \\
I \arrow[r, "x"] & J & S
\end{tikzcd}
$$

With this picture in mind, we shall call a morphism (x, f) of $\mathcal{G}(\mathbb{C})$ *vertical* if it lies in a fibre of Π (i.e. if x is an identity morphism), and *horizontal* if f is an identity morphism. As the above diagram indicates, every morphism of $\mathcal{G}(\mathbb{C})$ can be factored uniquely into a vertical morphism followed by a horizontal one.

Up to isomorphism, the horizontal morphisms can be described purely in terms of the functor $\Pi \colon \mathcal{G}(\mathbb{C}) \to S$:

Lemma 1.3.2 *Let $h \colon U \to V$ be a morphism of $\mathcal{G}(\mathbb{C})$. Then the vertical part of the factorization of h is an isomorphism iff, given any morphism $k \colon W \to V$ with the same codomain and any factorization $\Pi(k) = \Pi(h) \circ x$ of $\Pi(k)$ through $\Pi(h)$, there is a unique $l \colon W \to U$ in $\mathcal{G}(\mathbb{C})$ with $\Pi(l) = x$ and $hl = k$.*

Proof Suppose $h = (y, f)$ where f is an isomorphism, and let $k = (yx, g)$. Then the required factorization of k through h is easily seen to be $(x, x^*(f^{-1}) \circ g)$ (modulo composition with a coherence isomorphism whose name we have suppressed). Conversely, if $h = (y, f)$ satisfies the given condition, take $k = (y, 1)$ and $x = 1$; the second component of the unique vertical factorization of k through h must be a two-sided inverse for f. □

Now let $\Pi \colon C \to S$ be an arbitrary functor. Inspired by the foregoing result, we shall call a morphism of C Π-*prone* (or simply *prone*, if there is no need to specify Π) if it satisfies the condition in the statement of Lemma 1.3.2, and Π-*vertical* if its image under Π is an identity morphism. Note that this condition is not self-dual, which is our reason for abandoning the word 'horizontal'; a morphism which satisfies the dual condition (i.e. one which is prone

with respect to $\Pi^{\mathrm{op}}\colon \mathcal{C}^{\mathrm{op}} \to \mathcal{S}^{\mathrm{op}}$) will be called Π-*supine*. (Prone morphisms are more commonly called 'cartesian', but this word has been rather overworked by category-theorists, and deserves a rest. The terminology we have adopted, which conveniently combines the idea of lying horizontally with a 'handedness', was suggested by P. Taylor [1161].) We shall need the following properties of prone morphisms:

Lemma 1.3.3

(i) *Let $f\colon A \to B$ and $g\colon B \to C$ be morphisms of $\mathcal{C}$ such that g is prone. Then the composite gf is prone iff f is.*

(ii) *Given a commutative square*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"] & B \arrow[d, "h"] \\ C \arrow[r, "k"] & D \end{tikzcd}$$

in $\mathcal{C}$ such that k is prone and g and h are vertical, then f is prone iff the square is a pullback.

Proof Straightforward verification. □

By 1.3.3(i), the prone morphisms of $\mathcal{C}$ form a subcategory $\mathcal{C}_p$ (it is clear that they include all identity morphisms, and indeed all isomorphisms of $\mathcal{C}$); and for each object A of $\mathcal{C}$ the slice category $\mathcal{C}_p/A$ is a full subcategory of $\mathcal{C}/A$. We shall also write $\mathcal{C}^I$ for the *fibre* of $\mathcal{C}$ over an object I of $\mathcal{S}$, i.e. the subcategory of objects mapping to I and vertical morphisms between them.

Definition 1.3.4 A functor $\Pi\colon \mathcal{C} \to \mathcal{S}$ is called a *fibration* if, for any object A of $\mathcal{C}$ and any morphism $x\colon I \to \Pi(A)$ in $\mathcal{S}$, there exists a Π-prone morphism $f\colon B \to A$ in $\mathcal{C}$ with $\Pi(f) = x$.

We do not, in general, demand that f should be unique; but it will automatically be unique up to vertical isomorphism; that is, if two prone morphisms f and f' have the same codomain and the same image under Π, there is a unique vertical isomorphism v such that $f' = fv$.

If $\Pi\colon \mathcal{C} \to \mathcal{S}$ is a fibration, a *cleavage* for Π is a function σ assigning to each pair (A, x) with $\Pi(A) = \operatorname{cod} x$ a prone morphism $f = \sigma(A, x)$ with $\operatorname{cod} f = A$ and $\Pi(f) = x$. We shall usually assume that σ is *normalized* in the sense that $\sigma(A, 1_{\Pi(A)}) = 1_A$; if not, it is easy to modify σ to achieve this. We say σ is a

splitting if it satisfies the compatibility condition

$$\sigma(A, yx) = \sigma(A, y) \circ \sigma(\text{dom }\sigma(A, y), x)$$

for every composable pair of morphisms (x, y) in $\mathcal{S}$; in general, it is not possible to modify a given σ to achieve this. By a *cloven* (resp. *split*) *fibration* we mean a fibration equipped with a choice of a cleavage (resp. a splitting).

The notion of cloven fibration may be defined in an essentially algebraic way – that is, it may be specified by partial operations whose domains are defined by equations in the previously specified operations, and satisfying equations (cf. [371]) – and so it may be interpreted in the internal logic of any cartesian category. Specifically, the structure involved is that of a pair of categories $\mathcal{C}$ and $\mathcal{S}$, a functor $\Pi\colon \mathcal{C} \to \mathcal{S}$, the operation σ assigning to pairs (A, x) with $\Pi(A) = \text{cod }x$ a morphism $\sigma(A, x)$ of $\mathcal{C}$ with $\Pi(\sigma(A, x)) = x$ and $\text{cod }\sigma(A, x) = A$, and an operation τ assigning to a triple (x, y, f) with (x, y) a composable pair in $\mathcal{S}$ and $\Pi(f) = yx$ the unique morphism $\tau(x, y, f)$ satisfying $\Pi(\tau(x, y, f)) = x$ and $\sigma(\text{cod }f, y) \circ \tau(x, y, f) = f$. We leave to the reader the task of formulating the other equations which σ and τ must satisfy.

For the fibration $\Pi\colon \mathcal{G}(\mathbb{C}) \to \mathcal{S}$ arising from an $\mathcal{S}$-indexed category $\mathbb{C}$, we have a canonical choice of cleavage given by $\sigma(A, x) = (x, 1_{x^*A})$; this cleavage is a splitting iff $\mathbb{C}$ is a strict functor and not just a pseudofunctor. In the converse direction, we have

Theorem 1.3.5 *Let $\Pi\colon \mathcal{C} \to \mathcal{S}$ be a fibration, and σ a cleavage for Π. Then there is a pseudofunctor $\mathbb{C}\colon \mathcal{S}^{\text{op}} \to \mathfrak{CAT}$ sending an object I of $\mathcal{S}$ to $\mathcal{C}^I$, and such that the Grothendieck construction applied to this indexed category yields a fibration isomorphic (in $\mathfrak{CAT}/\mathcal{S}$) to Π. Moreover, $\mathbb{C}$ is a strict functor iff σ is a splitting.*

Proof More straightforward verification. Given a morphism $x\colon I \to J$ in $\mathcal{S}$, the induced functor $x^*\colon \mathcal{C}^J \to \mathcal{C}^I$ sends an object A of $\mathcal{C}^J$ to $\text{dom }\sigma(A, x)$, and a morphism $f\colon A \to B$ to the unique vertical factorization of $f \circ \sigma(A, x)$ through $\sigma(B, x)$. The functoriality of x^* follows from the uniqueness of such vertical factorizations. Given $y\colon J \to K$ and an object C of $\mathcal{C}^K$, we have a canonical vertical isomorphism $x^*y^*C \to (yx)^*C$ arising from the fact that both objects are the domains of prone morphisms to C; this defines the required natural isomorphism $x^*y^* \to (yx)^*$, and it is an identity (for all pairs (x, y)) iff σ is a splitting. Finally, there is an obvious isomorphism $\mathcal{G}(\mathbb{C}) \to \mathcal{C}$ which sends an object (I, A) to A and a morphism $(x, f)\colon (I, A) \to (J, B)$ to the composite $\sigma(B, x) \circ f$. □

Thus $\mathcal{S}$-indexed categories are 'the same thing as' cloven fibrations over $\mathcal{S}$, and we shall henceforth use 1.3.5 as a licence to pass back and forth between the two kinds of structure, using whichever is more convenient at any given moment. In order to do this, we need to make 1.3.5 functorial; that is, to investigate

the concepts for fibrations which correspond to $\mathcal{S}$-indexed functors and natural transformations.

Rather against our general philosophy, a morphism of fibrations from $\Pi\colon \mathcal{C} \to \mathcal{S}$ to $\Phi\colon \mathcal{D} \to \mathcal{S}$ will (in particular) be a functor $F\colon \mathcal{C} \to \mathcal{D}$ making the diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "F"] \arrow[dr, "\Pi"'] & \mathcal{D} \arrow[d, "\Phi"] \\ & \mathcal{S} \end{tikzcd}$$

commute strictly, and not just up to isomorphism. (This is not really as arbitrary as it seems; if we had an F making the above diagram commute up to a natural isomorphism $\alpha\colon \Pi \to \Phi F$, we could use a cleavage for Φ to 'lift' α to a natural isomorphism $F' \to F$, where F' is a functor making the diagram commute strictly.) For F to be a *morphism of fibrations,* we further require that it should send Π-prone morphisms to Φ-prone ones (the fact that the diagram commutes ensures that it sends Π-vertical morphisms to Φ-vertical ones); however, if Π and Φ are equipped with cleavages, we do not demand that F should commute with them in any sense. A *(vertical) transformation* between morphisms of fibrations F and G is of course a natural transformation $\alpha\colon F \to G$ which is compatible with Π and Φ (i.e. such that each α_A is Φ-vertical). We write $\mathfrak{Fib}_{\mathcal{S}}$ (resp. $\mathfrak{cFib}_{\mathcal{S}}$, $\mathfrak{sFib}_{\mathcal{S}}$) for the 2-category whose objects are fibrations (resp. cloven fibrations, split fibrations) over $\mathcal{S}$, with the above morphisms and transformations.

Theorem 1.3.6 *The Grothendieck construction yields an equivalence of 2-categories between $\mathfrak{CAT}_{\mathcal{S}}$ and $\mathfrak{cFib}_{\mathcal{S}}$.*

Proof Straightforward verification again. □

Examples 1.3.7 (*a*) Let $\mathcal{S}$ be a cartesian category. The fibration which corresponds to the $\mathcal{S}$-indexed category $\mathbb{S}$ of Example 1.2.2(*c*) is simply $\mathrm{cod}\colon [\mathbf{2}, \mathcal{S}] \to \mathcal{S}$, where $\mathbf{2}$ is the finite category represented diagrammatically by

$$\bullet \longrightarrow \bullet$$

(so that functors $\mathbf{2} \to \mathcal{S}$ may be identified with morphisms of $\mathcal{S}$, and natural transformations between them with commutative squares in $\mathcal{S}$) and cod is the functor sending a morphism of $\mathcal{S}$ to its codomain. It is clear that the fibre of this functor over I is exactly the slice category $\mathcal{S}/I$; and a commutative square in $\mathcal{S}$, regarded as a morphism of $[\mathbf{2}, \mathcal{S}]$, is prone relative to cod iff it is a pullback square, from which it follows that the transition functors between the fibres are indeed the pullback functors defined in A1.2.8. Indeed, it is not hard to see that, for an arbitrary $\mathcal{S}$, the functor $\mathrm{cod}\colon [\mathbf{2}, \mathcal{S}] \to \mathcal{S}$ is a fibration iff $\mathcal{S}$ has pullbacks – and that the difference between the two possible interpretations

of the statement '$\mathcal{S}$ has pullbacks', on which we touched after Lemma A1.2.1, is precisely the difference between merely asserting that cod is a fibration and asserting that it is a cloven fibration. We remark that cod is also an opfibration (that is, $\mathrm{cod}^{\mathrm{op}} \colon [\mathbf{2}, \mathcal{S}]^{\mathrm{op}} \to \mathcal{S}^{\mathrm{op}}$ is a fibration), and indeed a split one: supine morphisms of $[\mathbf{2}, \mathcal{S}]$ are those squares whose 'top edge' is an isomorphism, and the splitting sends (x, y) to the square

$$\begin{tikzcd} \bullet \arrow[r, "1"] \arrow[d, "x"] & \bullet \arrow[d, "yx"] \\ \bullet \arrow[r, "y"] & \bullet \end{tikzcd}$$

We shall see in 1.4.5 below that this property of cod corresponds to the fact that the transition functors x^* of the corresponding indexed category have left adjoints Σ_x (and that the latter are strictly functorial in x).

(*b*) More generally, if $F \colon \mathcal{S} \to \mathcal{C}$ is a functor between cartesian categories, then the fibration corresponding to the $\mathcal{S}$-indexed category $\mathbb{C}$ of 1.2.2(*d*) is simply the projection functor $P_2 \colon \mathbf{Gl}(F) \to \mathcal{S}$, where $\mathbf{Gl}(F)$ is the category obtained by glueing along F, as defined in A2.1.12.

(*c*) Again, the fibration corresponding to the indexed category $\mathbb{S}\mathrm{ub}$ of 1.2.2(*g*) is (the restriction of cod to) the full subcategory $\mathbf{Mono}(\mathcal{S})$ of $[\mathbf{2}, \mathcal{S}]$ whose objects are monomorphisms in $\mathcal{S}$.

(*d*) The fibration corresponding to the indexed category $\mathbb{T}$ of 1.2.2(*h*) is none other than the forgetful functor $\mathbf{Sp} \to \mathbf{Set}$; for the objects of its domain are simply sets equipped with a topology, and morphisms $(I, \mathcal{U}) \to (J, \mathcal{V})$ are functions $x \colon I \to J$ such that $\mathcal{U} \supseteq x^*\mathcal{V}$, i.e. such that x is continuous for the given topologies. Note that this functor is also an opfibration; again, this corresponds to the fact that the x^* have left adjoints, as we noted in 1.2.2(*h*).

Change of base for indexed categories, as defined in 1.2.2(*e*), corresponds to pullback of fibrations, in the following sense:

Lemma 1.3.8 *Let $F \colon \mathcal{S} \to \mathcal{T}$ be a functor, and $\mathbb{C}$ a $\mathcal{T}$-indexed category. Then there is a pullback square*

$$\begin{tikzcd} \mathcal{G}(F^*\mathbb{C}) \arrow[r] \arrow[d, "\Phi"] & \mathcal{G}(\mathbb{C}) \arrow[d, "\Pi"] \\ \mathcal{S} \arrow[r, "F"] & \mathcal{T} \end{tikzcd}$$

where the vertical arrows are fibrations.

Proof By definition, the objects of the pullback are pairs $(I,(J,A))$ where $I \in \text{ob}\,\mathcal{S}$, $(J,A) \in \text{ob}\,\mathcal{G}(\mathbb{C})$ and $FI = \Pi(J,A) = J$; but these correspond exactly to the objects (I,A) of $\mathcal{G}(F^*\mathbb{C})$. It is straightforward to verify that this correspondence defines an isomorphism of categories. □

Once again, it will be observed that the notion of pullback used in the above proof is the 'strict' one, and not the pseudo-limit as described in Section B1.1. However, for fibrations there is no essential difference:

Lemma 1.3.9 *Let $\Pi: \mathcal{C} \to \mathcal{T}$ be a cloven (resp. split) fibration, and $F: \mathcal{S} \to \mathcal{T}$ an arbitrary functor. Then the strict pullback of Π along F is a cloven (resp. split) fibration, and it is also a pseudo-pullback.*

Proof Let us denote the (strict) pullback by

$$\begin{tikzcd} \mathcal{D} \arrow[r, "G"] \arrow[d, "\Phi"] & \mathcal{C} \arrow[d, "\Pi"] \\ \mathcal{S} \arrow[r, "F"] & \mathcal{T} \end{tikzcd} .$$

Then, given an object A of $\mathcal{D}$ and $x: I \to \Phi A$ in $\mathcal{S}$, it is easily verified that $(x, \sigma(GA, Fx))$ is a prone morphism in $\mathcal{D}$ with codomain A for any cleavage σ of Π, and that the assignment $(A,x) \mapsto (x, \sigma(GA, Fx))$ is a cleavage of Φ (and a splitting if σ is). For the second assertion, we note that the usual construction of the pseudo-pullback of F and Π in $\mathfrak{CAT}$ would produce a category $\widetilde{\mathcal{D}}$ whose objects are triples (I,x,A) with $I \in \text{ob}\,\mathcal{S}$, $A \in \text{ob}\,\mathcal{C}$ and $x: FI \to \Pi A$ an isomorphism in $\mathcal{T}$. Clearly the assignment $(I,A) \mapsto (I, 1_{FI}, A)$ defines a functor $\mathcal{D} \to \widetilde{\mathcal{D}}$; and when Π is a cloven fibration we can construct an inverse for it up to natural isomorphism, sending (I,x,A) to $(I, \text{dom}(\sigma(x,A)))$. So the strict pullback is equivalent to the pseudo-pullback, and hence shares the universal property of the latter. □

Our convention that a cartesian category comes equipped with a choice of 'canonical' finite limits will ensure that virtually all the fibrations we have to consider will come equipped with cleavages, although many of them will not have splittings. However, if we are simply given a fibration (or even an arbitrary functor), we can always replace it by a 'best approximation' amongst split fibrations, by the following device:

Definition 1.3.10 Let $\Pi: \mathcal{C} \to \mathcal{S}$ be an arbitrary functor. The *associated split fibration* of Π is the functor $\hat{\Pi}: \hat{\mathcal{C}} \to \mathcal{S}$, where an object of $\hat{\mathcal{C}}$ consists of an object A of $\mathcal{C}$ together with a functor $\sigma: \mathcal{S}/\Pi A \to \mathcal{C}/A$ such that every $\sigma(x)$ is Π-prone and satisfies $\Pi(\sigma(x)) = x$. (As usual, we also demand that $\sigma(1_{\Pi A}) = 1_A$.) A morphism $(A,\sigma) \to (A',\sigma')$ in $\hat{\mathcal{C}}$ consists of a morphism $x: \Pi A \to \Pi A'$ in $\mathcal{S}$

together with a morphism $f\colon A \to \operatorname{dom}\sigma'(x)$ in $\mathcal{C}^{\Pi A}$ (which of course induces, for each y with codomain ΠA, a unique morphism $\operatorname{dom}\sigma(y) \to \operatorname{dom}\sigma'(xy)$ in $\mathcal{C}^{\operatorname{dom} y}$). Composition and identity morphisms in $\hat{\mathcal{C}}$ are defined in the obvious way; the functor $\hat{\Pi}$ sends (A, σ) to ΠA and (x, f) to x.

$\hat{\Pi}$ is a split fibration: to define the splitting at an object (A, σ) and a morphism $x\colon I \to \Pi A$, we take its domain to be $(\operatorname{dom}\sigma(x), \sigma')$ where $\sigma'(y) = \sigma(y)$ for all $y\colon J \to I$ (and y is considered on the left of the latter equation as an object of $\mathcal{S}/I$, and on the right as a morphism $xy \to x$ in $\mathcal{S}/\Pi A$). There is a factorization $U\colon \hat{\mathcal{C}} \to \mathcal{C}$ of $\hat{\Pi}$ through Π, sending (A, σ) to A and (x, f) to $\sigma'(x) \circ f$; this functor is always full and faithful. If Π is a cloven fibration with cleavage σ, then we can define an inverse (up to isomorphism) for U, sending an object A of $\mathcal{C}$ to $(A, \sigma(A, -))$. Thus we have shown that *every cloven fibration over $\mathcal{S}$ is equivalent (in $\mathfrak{CAT}/\mathcal{S}$) to a split fibration.*

There is an important special case in which no arbitrary choices are required to obtain a splitting of a fibration. We shall say a category is *stiff* if its only isomorphisms are identity morphisms; and that a fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$ is *stiff* if its fibres $\mathcal{C}^I$ are all stiff categories, i.e. the only vertical isomorphisms in $\mathcal{C}$ are identities. (Stiffness is the conjunction of two more familiar categorical properties usually called 'skeletal' and 'rigid'.) If Π has this property, then any morphism $x\colon I \to J$ of $\mathcal{S}$ has a unique prone lifting with codomain A, for each $A \in \operatorname{ob}\mathcal{C}^J$; so Π has a unique cleavage, which is in fact a splitting. As a further specialization, we may consider *discrete fibrations*, which are those in which every fibre $\mathcal{C}^I$ is a discrete category. An important property of discrete fibrations is worth recording:

Lemma 1.3.11 *Let $\Pi\colon \mathcal{C} \to \mathcal{S}$ and $\Phi\colon \mathcal{D} \to \mathcal{S}$ be fibrations, Φ being discrete, and let $F\colon \mathcal{C} \to \mathcal{D}$ be a functor satisfying $\Phi F = \Pi$. Then*

(i) *F is a morphism of fibrations.*
(ii) *F is itself a fibration.*
(iii) *F is a discrete fibration iff Π is.*

Proof (i) Every morphism of $\mathcal{D}$ is Φ-prone, for in its vertical–prone factorization the vertical part must be an identity morphism. So this is immediate from the definition of a morphism of fibrations.

(ii) First we claim that a morphism $f\colon A \to B$ of $\mathcal{C}$ is Π-prone iff it is F-prone. For suppose f is Π-prone and we are given $h\colon C \to B$ and a factorization $F(h) = F(f)g$ in $\mathcal{D}$; then $\Pi(h) = \Pi(f)\Phi(g)$ so that we have a unique $k\colon C \to A$ with $fk = h$ and $\Pi(k) = \Phi(g)$, and then $F(k) = g$ since both are (necessarily prone) liftings of $\Phi(g)$ with codomain $F(A)$. Conversely, if f is F-prone, then since $F(f)$ is Φ-prone it follows easily that f is Π-prone.

Now, given an object A of $\mathcal{C}$ and a morphism $f\colon B \to F(A)$ in $\mathcal{D}$, we have a Π-prone $g\colon C \to A$ with $\Pi(g) = \Phi(f)$, and then $F(g) = f$ since both are liftings of $\Phi(f)$ with codomain $F(A)$. So g is an F-prone lifting of f.

(iii) Given (ii), this is almost immediate from the definition of a discrete fibration. □

Given the fact that discrete fibrations over a given category $\mathcal{S}$ with small fibres correspond to (necessarily strict) functors $\mathcal{S}^{\mathrm{op}} \to \mathbf{Set}$, and that morphisms of fibrations between them correspond to natural transformations, the conjunction of (i) and (iii) of Lemma 1.3.11 is simply a reformulation (for contravariant functors, this time) of Proposition A1.1.7.

An important concept in the theory of fibrations (or of indexed categories) is that of *comprehensibility* (and the special case of it known as *definability*). This concept is so widespread that it is hard to give a simple definition of it which will cover all cases of interest. However, the basic idea is as follows: suppose given a construction which, when applied to some finite collection of data is an ordinary category $\mathcal{C}$, would produce a class (or if $\mathcal{C}$ were 'small enough', a set) of 'structures' on the given data. Then, if we apply the construction to data in the total category of a fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$, we should expect to obtain not just a single class but a contravariant functor defined on $\mathcal{S}/I_1 \times \cdots \times I_n$ (where the given data lie in the fibres $\mathcal{C}^{I_1}, \ldots, \mathcal{C}^{I_n}$); and we say that the result of the construction is *comprehensible* in $\mathcal{C}$ if this functor is representable (for all choices of the given data).

To make this more precise, we shall interpret the notion of 'collection of data' as being given by a diagram of some (fixed) finite shape, and 'structure' as being given by an extension of the diagram to a larger one. Given a fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$ and a finite category $\mathcal{D}$, we define the *rectangular diagram category* $\mathrm{Rect}(\mathcal{D}, \mathcal{C})$ to be the non-full subcategory of $[\mathcal{D}, \mathcal{C}]$ whose objects are the 'vertical diagrams' of shape $\mathcal{D}$ (i.e. diagrams with all edges vertical) and whose morphisms are natural transformations with all components prone. Then we say that $\Pi\colon \mathcal{C} \to \mathcal{S}$ *satisfies the comprehension scheme* for a functor $G\colon \mathcal{D}' \to \mathcal{D}$ if the functor $\mathrm{Rect}(\mathcal{D}, \mathcal{C}) \to \mathrm{Rect}(\mathcal{D}', \mathcal{C})$ obtained by composition with G has a right adjoint.

In order to understand this rather opaque definition, it is best to consider a few examples.

Example 1.3.12 We say that a fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$ (or the corresponding $\mathcal{S}$-indexed category) is *locally small* if it satisfies the comprehension scheme for the inclusion functor $2 \to \mathbf{2}$, where 2 is the discrete category with two objects and $\mathbf{2}$ is the category $(\bullet \to \bullet)$. In elementary terms, this means the following: given any two objects A and B of $\mathcal{C}$, there exists an object $(x, y)\colon I \to \Pi A \times \Pi B$ of $\mathcal{S}/\Pi A \times \Pi B$, and a morphism $f\colon x^*A \to y^*B$ in the fibre $\mathcal{C}^I$ which is *generic* in the sense that, given any $(z, w)\colon J \to \Pi A \times \Pi B$ and any $g\colon z^*A \to w^*B$ in $\mathcal{C}^J$, there is a unique $u\colon J \to I$ such that $xu = z$, $yu = w$ and u^*f is identified (modulo the coherence isomorphisms) with g.

If $\mathcal{C}$ is any locally small category (in the 'classical' sense), then the naive indexing of $\mathcal{C}$ over $\mathbf{Set}$ is a locally small indexed category: given objects $A = (A_j \mid j \in J)$ and $B = (B_k \mid k \in K)$, we take I to be the disjoint union, over all

pairs (j, k), of the hom-sets $\mathcal{C}(A_j, B_k)$, with the obvious projection to $J \times K$. (It is easy to see that the converse is true, by considering objects in the fibre over 1.) However, local smallness of the fibres in the classical sense is not a sufficient condition for local smallness of a fibration over **Set**: for example, the fibration $\mathbf{Sp} \to \mathbf{Set}$ of 1.3.7(*d*) is not locally small, even though its fibres are preorders. (To see this, let $\mathcal{U}$ and $\mathcal{V}$ be topologies on sets J and K respectively; then since there is a unique topology on a singleton set we see that any morphism $(j, k): 1 \to J \times K$ would have to factor uniquely through the morphism $I \to J \times K$ of 1.3.12 if it existed, and hence the latter would have to be bijective. But this would imply that we had $\pi_1^*\mathcal{U} \supseteq \pi_2^*\mathcal{V}$ as topologies on $J \times K$, which is clearly not true in general.)

It is also easy to see that the discrete fibration corresponding to the indexed category $\mathbb{I}$ of Example 1.2.5 is locally small iff $\mathcal{S}$ has pullbacks over I (i.e. the slice category $\mathcal{S}/I$ has finite products), since the object $P \to J \times K$ representing vertical morphisms from $x: J \to I$ to $y: K \to I$ must be exactly a pullback of x against y. We shall investigate local smallness for indexed categories in greater detail in Section B2.2; in particular, we shall see there that the canonical indexing $\mathbb{S}$ of a cartesian category $\mathcal{S}$ over itself is locally small iff $\mathcal{S}$ is locally cartesian closed.

The comprehension scheme works best when (as in the above example) the functor $G: \mathcal{D}' \to \mathcal{D}$ is bijective on objects. If we apply it to a functor without this property, we can obtain surprising results, as the following example shows.

Example 1.3.13 One might be tempted to call a fibration $\Pi: \mathcal{C} \to \mathcal{S}$ *small* if it (is locally small and) satisfies the comprehension scheme for the inclusion $0 \to 1$, where 0 is the empty category and 1 the singleton category. In elementary terms, this means that the subcategory $\mathcal{C}_p$ of prone morphisms of $\mathcal{C}$ has a terminal object. However, although this condition is of use in certain contexts, when considered as a definition of smallness it says both rather less and rather more than one might wish. On the one hand, it asserts the representability of the functor which to an object I of $\mathcal{S}$ assigns the collection of isomorphism classes of objects of $\mathcal{C}^I$, rather than ob $\mathcal{C}^I$ itself; thus it is really a notion of 'essential smallness'. (Of course, for a fibration which does not possess a splitting, the assignment $I \mapsto$ ob $\mathcal{C}^I$ cannot be made into a functor on $\mathcal{S}^{\mathrm{op}}$, so it does not make sense to ask that it should be representable.) But it also implies that the fibres of $\mathcal{C}$ are rigid, i.e. that any vertical automorphism in $\mathcal{C}$ is an identity morphism; for if an object A admits a vertical automorphism f and a prone morphism $g: A \to B$, then gf is also prone, and the equality $gf = g$ forces $f = 1_A$. In Section B2.3 below, we shall develop a more satisfactory notion of 'small $\mathcal{S}$-indexed category' from an entirely different starting-point.

In certain cases, we shall wish to consider a 'generalized comprehension scheme' where we seek a right adjoint not for $\mathrm{Rect}(\mathcal{D}, \mathcal{C}) \to \mathrm{Rect}(\mathcal{D}', \mathcal{C})$ but for the restriction of this functor to some full subcategory of its domain, defined by requiring the objects to be diagrams having some additional property (e.g.

that certain edges should be monic, or that certain parts of the diagrams should be limits). Of course, in this case we must restrict our attention to fibrations whose re-indexing functors preserve the given properties of diagrams. A good example is the following:

Example 1.3.14 Let $\Pi: \mathcal{C} \to \mathcal{S}$ be a fibration such that the transition functors $x^*: \mathcal{C}^I \to \mathcal{C}^J$ induced by morphisms $x: J \to I$ of $\mathcal{S}$ preserve monomorphisms. (Note that this will be the case if the x^* have left adjoints – a condition which we shall explore more fully in the next section.) We say that Π (or the corresponding $\mathcal{S}$-indexed category) is *well-powered* if the forgetful functor $U: Q(\mathbf{2}, \mathcal{C}) \to \mathrm{Rect}(1, \mathcal{C})$ has a right adjoint, where $Q(\mathbf{2}, \mathcal{C})$ is the full subcategory of $\mathrm{Rect}(\mathbf{2}, \mathcal{C})$ whose objects are vertical monomorphisms and U sends each monomorphism to its codomain. In elementary terms, this means that for each $A \in \mathrm{ob}\, \mathcal{C}$ we have a morphism $x: I \to \Pi A$ in $\mathcal{S}$ and a subobject $A' \rightarrowtail x^*A$ in $\mathcal{C}^I$ such that, for any $y: J \to \Pi A$ and any $A'' \rightarrowtail y^*A$ in $\mathcal{C}^J$, there is a unique $z: J \to I$ such that $xz = y$ and $z^*(A') \cong A''$ in $\mathrm{Sub}_{\mathcal{C}^J}(y^*A)$. (Note that, once again, the fact that we are dealing with (possibly) non-split fibrations forces us to interpret the word 'subobject' here as meaning an isomorphism class of monomorphisms with given codomain, and not an individual monomorphism; but in the present context this is what we should wish to do anyway. Note also that the enforced rigidity, which we observed in 1.3.13, does not bother us here, since categories of the form $\mathrm{Sub}(A)$ are (preorders, and hence) rigid in any case.)

It is clear that the naive indexing of an ordinary category $\mathcal{C}$ over **Set** yields a well-powered fibration iff $\mathcal{C}$ is well-powered in the classical sense, as defined in Section A1.4. More interestingly, the canonical indexing $\mathbb{S}$ of a cartesian category $\mathcal{S}$ over itself is well-powered iff each of the slice categories $\mathcal{S}/I$ have power objects as defined in Section A2.1; but by A2.3.2 this is equivalent to saying that $\mathcal{S}$ is a topos.

For future reference, we note that the canonical indexing of a topos $\mathcal{S}$ over itself is also well-copowered (that is, $\mathbb{S}^{\mathrm{op}}$ is well-powered). Since a topos is effective regular (and since all epimorphisms in it are regular), we know that quotients of a given object I of $\mathcal{S}$ correspond bijectively to equivalence relations on I, in the sense of A1.3.6; and the latter can be 'classified' by an appropriate subobject $E(I)$ of $P(I \times I)$. Moreover, this construction works in any slice category $\mathcal{S}/J$; and it is preserved by pullback functors since they are logical. So we deduce that $\mathcal{S}/E(I)$ contains a generic quotient of I, in the sense required by the definition.

Definability is the special case of comprehensibility where we are concerned with properties of a collection of data, rather than constructions on it (or equivalently, with constructions that produce either a singleton or nothing at all). Typically, this occurs when the functor G along which we are trying to extend diagrams is an epimorphism in **Cat**. Two examples, which will be of importance in Section B2.4, are definability of equality for parallel pairs of morphisms, and definability of invertibility for morphisms. The first is simply the comprehension

scheme for $G: \mathcal{P} \to \mathbf{2}$, where $\mathcal{P}$ is the parallel-pair category ($\bullet \rightrightarrows \bullet$) and G identifies the two non-identity morphisms of $\mathcal{P}$, and the second is the comprehension scheme for the inclusion $\mathbf{2} \to \mathcal{I}$, where $\mathcal{I}$ is the trivial connected groupoid with two objects.

Lemma 1.3.15 *Let $\mathcal{S}$ be a category with equalizers. Then any locally small fibration $\Pi: \mathcal{C} \to \mathcal{S}$ has definable equality. If $\mathcal{S}$ also has pullbacks, then any such $\mathcal{C}$ also has definable invertibility.*

Proof (i) Given $(f,g): A \rightrightarrows B$ in $\mathcal{C}^I$ for some I, let $J \to I \times I$ be the object of $\mathcal{S}/I \times I$ indexing morphisms from A to B. Then f and g correspond to morphisms $I \rightrightarrows J$ in $\mathcal{S}$, and we have simply to take the equalizer of this pair to obtain a subobject $I' \rightarrowtail I$ such that an arbitrary $x: K \to I$ factors through I' iff $x^*f = x^*g$. But this is precisely what we seek.

(ii) Now consider a single morphism $f: A \to B$ in $\mathcal{C}^I$. Let $x: J \to I$ be the pullback along the diagonal $I \rightarrowtail I \times I$ of the object indexing vertical morphisms $B \to A$ in $\mathcal{C}$, and let $g: x^*(B) \to x^*(A)$ be the generic such morphism in $\mathcal{C}^J$. If we form the intersection K of the two subobjects of J measuring the equality of the pairs $(x^*(f) \circ g, 1_{x^*(B)})$ and $(g \circ x^*(f), 1_{x^*(A)})$ in $\mathcal{C}^J$, then factorizations of an arbitrary $y: L \to I$ through the composite $K \rightarrowtail J \to I$ correspond bijectively to inverses for $y^*(f)$ – but since $y^*(f)$ has at most one inverse for any given y, it follows that this composite must be a monomorphism, and it is the subobject of I which we seek. □

Generalizing the proof of 1.3.15, it can be shown that if $\mathcal{S}$ is cartesian and $\Pi: \mathcal{C} \to \mathcal{S}$ is a locally small fibration, then Π satisfies the comprehension scheme for any $G: \mathcal{D}' \to \mathcal{D}$ which is bijective on objects.

We shall also be interested in definability of subfibrations of a given fibration. Of course, if $\mathcal{C}$ is equipped with a fibration $\Pi: \mathcal{C} \to \mathcal{S}$, we say that a subcategory $\mathcal{C}'$ of $\mathcal{C}$ is a *subfibration* if, whenever a prone morphism of $\mathcal{C}$ has its codomain in $\mathcal{C}'$, then the whole morphism belongs to $\mathcal{C}'$; this clearly implies (and, for replete subcategories, is equivalent to saying) that the restriction of Π to $\mathcal{C}'$ is still a fibration and that the inclusion $\mathcal{C}' \to \mathcal{C}$ is a morphism of fibrations. We say that a subfibration $\mathcal{C}'$ is definable in $\mathcal{C}$ if the inclusion $\mathrm{Rect}(\mathbf{2}, \mathcal{C}') \to \mathrm{Rect}(\mathbf{2}, \mathcal{C})$ has a right adjoint. (Note that this inclusion is full even if $\mathcal{C}' \to \mathcal{C}$ is not; so this is a generalized comprehension scheme as we defined it earlier.) In elementary terms, it means that for every $f: A \to B$ in $\mathcal{C}^I$, we can find a subobject $I' \rightarrowtail I$ such that an arbitrary $x: J \to I$ factors through I' iff x^*f lies in $\mathcal{C}'^J$.

For example, if an ordinary category $\mathcal{C}$ is indexed over **Set** in the naive way, as in 1.2.2(a), then the definable subfibrations of the corresponding fibration are just (the fibrations corresponding to the naive indexings of) the subcategories of $\mathcal{C}$: if $\mathcal{C}'$ is such a subcategory, and $(f_i \mid i \in I)$ is a morphism of $\mathcal{C}^I$, then the subset $I' \subseteq I$ which measures the extent to which this morphism is in $\mathcal{C}'$ is just $\{i \in I \mid f_i \in \text{mor } \mathcal{C}'\}$. But this fibration may have other full subfibrations which are not definable, for example that described (for $\mathcal{C} = \mathbf{Set}_f$) in 1.2.2($b$).

Slightly less trivially, we note that if $\mathcal{S}$ is a Heyting category, then (the fibration corresponding to) the $\mathcal{S}$-indexed category $\mathbb{S}$ub of 1.2.2(*g*) is definable as a full subfibration of that corresponding to the canonical indexing $\mathbb{S}$ of 1.2.2(c). To see this, let $x\colon K \to J$ be an arbitrary morphism of $\mathcal{S}$. We have to show that there is a subobject $y\colon I \rightarrowtail J$ such that, for any $z\colon L \to J$, $z^*(x)$ is monic iff z factors through y. To this end, form the kernel-pair $(a,b)\colon R \rightrightarrows K$ of x, and the diagonal morphism $d\colon K \rightarrowtail R$; now take y to be $\forall_{xa}(d)$. Then we see that z factors through y iff $z^*(d)$ is an isomorphism; but since z^* preserves kernel-pairs and equalizers this is equivalent to saying that the two morphisms forming the kernel-pair of $z^*(x)$ are equal, i.e. that $z^*(x)$ is monic.

Lemma 1.3.16

(i) *If $\mathbb{C}'$ is a definable subfibration of $\mathbb{C}$, then the inclusion* $\mathrm{Rect}(1,\mathbb{C}') \to \mathrm{Rect}(1,\mathbb{C})$ *has a right adjoint.*

(ii) *Suppose $\mathcal{S}$ has pullbacks. If $\mathbb{C}'$ is a full subfibration of $\mathbb{C}$ such that* $\mathrm{Rect}(1,\mathbb{C}') \to \mathrm{Rect}(1,\mathbb{C})$ *has a right adjoint, then $\mathbb{C}'$ is definable.*

(iii) *Suppose $\mathcal{S}$ has pullbacks and $\Pi\colon \mathbb{C} \to \mathcal{S}$ is locally small. Then a subfibration $\mathbb{C}'$ of $\mathbb{C}$ is definable iff (a) $\mathbb{C}'$ is locally small and (b) the full subfibration $\mathbb{C}''$ of $\mathbb{C}$ with the same objects as $\mathbb{C}'$ is definable.*

Proof (i) We may obtain this right adjoint by restricting the right adjoint of $\mathrm{Rect}(\mathbf{2},\mathbb{C}') \to \mathrm{Rect}(\mathbf{2},\mathbb{C})$ to the full subcategory of $\mathrm{Rect}(\mathbf{2},\mathbb{C})$ whose objects are identity morphisms in $\mathbb{C}$.

(ii) Conversely, if $\mathbb{C}'$ is full then a (vertical) morphism of $\mathbb{C}$ belongs to it iff its domain and codomain both do so. So, given a morphism $f\colon A \to B$ in $\mathbb{C}^I$, we form the intersection of the subobjects $I' \rightarrowtail I$ and $I'' \rightarrowtail I$ which 'measure the extent to which A and B are in $\mathbb{C}'$', and this yields the subobject which measures the extent to which f is in $\mathbb{C}'$.

(iii) If $\mathbb{C}'$ is definable, then the definability of $\mathbb{C}''$ follows from (i) and (ii), since $\mathrm{Rect}(1,\mathbb{C}'')$ coincides with $\mathrm{Rect}(1,\mathbb{C}')$. Also, if A and B are any two objects of $\mathbb{C}'$, let $(x,y)\colon I \to \Pi A \times \Pi B$ be the object indexing vertical morphisms from A to B in $\mathbb{C}$, and let $u\colon I' \rightarrowtail I$ measure the extent to which the generic vertical morphism $f\colon x^*A \to y^*B$ is in $\mathbb{C}'$. Then it is clear that $(xu, yu)\colon I' \to \Pi A \times \Pi B$ indexes morphisms from A to B in $\mathbb{C}'$, so $\mathbb{C}'$ is locally small.

Conversely, suppose $\mathbb{C}''$ is definable and $\mathbb{C}'$ is locally small. Given a morphism $f\colon A \to B$ in $\mathbb{C}^I$, we can find $u\colon I' \rightarrowtail I$ measuring the extent to which A and B both belong to $\mathbb{C}'$, as in (ii). The morphism u^*f may not lie in $\mathbb{C}'$, but it corresponds to a commutative diagram

$$\begin{tikzcd} I' \arrow[r, "\ulcorner u^*f \urcorner"] \arrow[dr, "\Delta"'] & J \arrow[d, "{(x,y)}"] \\ & I' \times I' \end{tikzcd}$$

where (x, y) is the object indexing morphisms from u^*A to u^*B in $\mathcal{C}$. Now let $(x', y')\colon J' \to I' \times I'$ be the object indexing morphisms from u^*A to u^*B in $\mathcal{C}'$; then the generic morphism $x'^*u^*A \to y'^*u^*B$ defines a morphism $J' \to J$ in $\mathcal{S}$, which is easily seen to be monic. Pulling back this monomorphism along $\ulcorner u^*f \urcorner$, we obtain a subobject I'' of I'; and then the composite monomorphism $I'' \rightarrowtail I' \rightarrowtail I$ measures the extent to which f lies in $\mathcal{C}'$. So $\mathcal{C}'$ is definable in $\mathcal{C}$. □

We note, in particular, that a locally small fibration may have subfibrations which are not locally small.

A question of considerable interest to us is when comprehensibility properties are preserved under change of base. Despite appearances to the contrary, the following proof does not require the base categories $\mathcal{S}$ and $\mathcal{T}$ to have finite products, but only pullbacks. However, we shall temporarily use the notation $\mathcal{S}/I_1 \times \cdots \times I_n$ (where $I_1, \ldots, I_n$ are objects of $\mathcal{S}$ for the category whose objects are objects J of $\mathcal{S}$ equipped with an n-tuple of morphisms $(x_p\colon J \to I_p \mid 1 \leq p \leq n)$; of course, if $\mathcal{S}$ does have products, this category is isomorphic to the slice category usually denoted thus. Note also that if we are given an n-tuple of morphisms $(y_p\colon I_p \to K_p \mid 1 \leq p \leq n)$, then the functor $\mathcal{S}/I_1 \times \cdots \times I_n \to \mathcal{S}/K_1 \times \cdots \times K_n$ obtained by composing with the y_p still has a right adjoint, obtained by pulling back consecutively along each of the y_p (in any order).

Proposition 1.3.17 *Let $\mathcal{S}$ and $\mathcal{T}$ be categories with pullbacks, $F\colon \mathcal{S} \to \mathcal{T}$ a functor having a right adjoint R, and $\Pi\colon \mathcal{C} \to \mathcal{T}$ a fibration. Then $F^*\Pi\colon F^*\mathcal{C} \to \mathcal{S}$ satisfies any (generalized) comprehension scheme satisfied by* Π.

Proof For simplicity, we shall deal only with the original case of a right adjoint for the functor $\mathrm{Rect}(\mathcal{D}, \mathcal{C}) \to \mathrm{Rect}(\mathcal{D}', \mathcal{C})$ induced by a functor $G\colon \mathcal{D}' \to \mathcal{D}$. The extension to appropriate full subcategories of $\mathrm{Rect}(\mathcal{D}, \mathcal{C})$ is an easy exercise.

Let $\pi_0\mathcal{D}$ denote the set of connected components of $\mathcal{D}$: then it is easily seen that applying Π to objects and morphisms of $\mathrm{Rect}(\mathcal{D}, \mathcal{C})$ yields a functor $\mathrm{Rect}(\mathcal{D}, \mathcal{C}) \to \mathcal{T}^{\pi_0\mathcal{D}}$. Hence we may construct a diagram

$$\begin{tikzcd}
\mathrm{Rect}(\mathcal{D}, F^*\mathcal{C}) \arrow[rr] \arrow[dr] \arrow[dd] & & \mathrm{Rect}(\mathcal{D}', F^*\mathcal{C}) \arrow[dr] \arrow[dd] & \\
& \mathrm{Rect}(\mathcal{D}, \mathcal{C}) \arrow[rr] \arrow[dd] & & \mathrm{Rect}(\mathcal{D}', \mathcal{C}) \arrow[dd] \\
\mathcal{S}^{\pi_0\mathcal{D}} \arrow[rr] \arrow[dr] & & \mathcal{S}^{\pi_0\mathcal{D}'} \arrow[dr] & \\
& \mathcal{T}^{\pi_0\mathcal{D}} \arrow[rr] & & \mathcal{T}^{\pi_0\mathcal{D}'}
\end{tikzcd}$$

in which all faces commute, and the left and right vertical ones are pullbacks.

From now on, we shall simplify further by assuming that $\mathcal{D}$ is connected, i.e. that $\pi_0\mathcal{D}$ is a singleton. (If not, we may deal separately with each component of $\mathcal{D}$; note that G maps each component of $\mathcal{D}'$ into a single component of $\mathcal{D}$.) However, we allow $\mathcal{D}'$ to have n components ($n \geq 0$). Thus an object $A\colon \mathcal{D}' \to F^*\mathcal{C}$ of $\mathrm{Rect}(\mathcal{D}', F^*\mathcal{C})$ lies over an n-tuple of objects $(I_1, \ldots, I_n)$ of $\mathcal{S}$; and its image FA in $\mathrm{Rect}(\mathcal{D}', \mathcal{C})$ lies over $(FI_1, \ldots, FI_n)$. Applying the right adjoint of $\mathrm{Rect}(\mathcal{D}, \mathcal{C}) \to \mathrm{Rect}(\mathcal{D}', \mathcal{C})$ to FA, we obtain an object $(y_p\colon J \to FI_p \mid 1 \leq p \leq n)$ of $\mathcal{T}/FI_1 \times \cdots \times FI_n$ and a G-extension B of $(y_1, \ldots, y_n)^*(FA)$ (that is, a diagram of shape $\mathcal{D}$ in $\mathcal{C}^J$ whose restriction along G is the diagram obtained by applying y_p^* to the pth component of FA, for each p) which is universal in the appropriate sense.

Now, by a straightforward extension of 1.2.3, the functor $\mathcal{S}/I_1 \times \cdots \times I_n \to \mathcal{T}/FI_1 \times \cdots \times FI_n$ induced by F has a right adjoint, obtained by applying R and then pulling back along $(\eta_{I_p}\colon I_p \to RFI_p \mid 1 \leq p \leq n)$ where η is the unit of $(F \dashv R)$. Applying this to $(y_p \mid 1 \leq p \leq n)$, we obtain an object $(x_p\colon K \to I_p \mid 1 \leq p \leq n)$ of $\mathcal{S}/I_1 \times \cdots \times I_n$ together with a (counit) morphism $z\colon FK \to J$ in $\mathcal{T}/FI_1 \times \cdots \times FI_n$. So if we apply z^* to the diagram B, we obtain a G-extension of $(y_1 z, \ldots, y_n z)^*(FA)$ in $\mathcal{C}^{FK}$, or equivalently of $(x_1, \ldots, x_n)^*(A)$ in $(F^*\mathcal{C})^K$. And since, for any object $(w_p\colon L \to I_p \mid 1 \leq p \leq n)$ of $\mathcal{S}/I_1 \times \cdots \times I_n$, morphisms $L \to K$ in $\mathcal{S}/I_1 \times \cdots \times I_n$ correspond to morphisms $FL \to J$ in $\mathcal{T}/FI_1 \times \cdots \times FI_n$, it is easy to see that this extension has the required universal property for the value at A of a right adjoint to $\mathrm{Rect}(\mathcal{D}, F^*\mathcal{C}) \to \mathrm{Rect}(\mathcal{D}', F^*\mathcal{C})$. □

Corollary 1.3.18 *Let $\mathcal{S}$ be (properly) cartesian closed, let I be an object of $\mathcal{S}$ and let $\mathbb{C}$ be an $\mathcal{S}$-indexed category. Then (the fibration corresponding to) the $\mathcal{S}$-indexed category $\mathbb{C}^I$ of* 1.2.2(f) *satisfies any comprehension scheme satisfied by $\mathbb{C}$.* □

In 2.2.6 below we shall obtain a partial converse to 1.3.18: if $\mathcal{T}$ is locally cartesian closed, then change of base along $F\colon \mathcal{S} \to \mathcal{T}$ preserves local smallness iff F has a right adjoint.

Suggestions for further reading: Bénabou [101, 102, 105].

B1.4 Limits and colimits

An important part of the reason for introducing the machinery of $\mathcal{S}$-indexed categories and fibrations is the desire to be able to talk about 'infinite' limits and colimits in elementary terms. We saw in Section A1.2 that finite limits (and colimits) are an elementary concept, and we can therefore extend these notions to indexed categories in a straightforward way: thus we say that an $\mathcal{S}$-indexed category $\mathbb{C}$ *has $\mathcal{S}$-indexed finite limits* (or is *$\mathcal{S}$-cartesian*) if each of the fibres

$\mathcal{C}^I$ has finite limits, and each transition functor $x^*: \mathcal{C}^J \to \mathcal{C}^I$ preserves them. (Similarly for finite limits of some particular type, such as equalizers. Note that the assertion that $\mathbb{C}$ has $\mathcal{S}$-indexed limits of some particular (finite) shape $\mathcal{D}$ is just the assertion that the indexed diagonal functor $\mathbb{C} \to [\![\mathcal{D}, \mathbb{C}]\!]$ has an indexed right adjoint, where $[\![\mathcal{D}, \mathbb{C}]\!]$ is the indexed functor category defined after 1.2.5.)

The possession of $\mathcal{S}$-indexed finite limits may also be neatly characterized in terms of fibrations:

Lemma 1.4.1 *Let $\mathcal{S}$ be a cartesian category, and $\Pi: \mathcal{C} \to \mathcal{S}$ a fibration. Then the corresponding $\mathcal{S}$-indexed category $\mathbb{C}$ is $\mathcal{S}$-cartesian iff $\mathcal{C}$ has finite limits and Π preserves them.*

Proof First suppose $\mathcal{C}$ has and Π preserves finite limits. Then it is clear that each fibre $\mathcal{C}^I$ of Π has finite connected limits (in particular, pullbacks) which are preserved by the inclusion $\mathcal{C}^I \to \mathcal{C}$, since the limit in $\mathcal{S}$ of a constant connected diagram on the object I is I itself. Moreover, it follows readily from 1.3.3(ii) that the transition functors $x^*: \mathcal{C}^I \to \mathcal{C}^J$ preserve pullbacks. So we need only verify that the fibres have terminal objects, which are preserved by the transition functors; but this is simply a matter of definition – if we define 1_I to be the domain of the prone morphism to the terminal object of $\mathcal{C}$ lying over $I \to 1$ in $\mathcal{S}$, then it is immediate from the definition of prone morphisms that 1_I is terminal in $\mathcal{C}^I$.

Conversely, suppose $\mathbb{C}$ is $\mathcal{S}$-cartesian. Given a finite diagram $D: \mathcal{J} \to \mathcal{C}$, we first form a limit cone (with summit I, say) over the diagram ΠD in $\mathcal{S}$, and then use the transition functors to pull back the vertices of the original diagram to the fibre $\mathcal{C}^I$. The definition of prone morphisms ensures that the edges of the original diagram pull back to vertical morphisms forming a diagram of shape $\mathcal{J}$ in $\mathcal{C}^I$; and if we form a limit cone over this diagram, it is again easy to see that its summit yields a limit for the original diagram in $\mathcal{C}$. Thus $\mathcal{C}$ has limits of shape $\mathcal{J}$, and Π preserves them. □

We shall use the term *cartesian fibration* for a fibration satisfying the equivalent conditions of 1.4.1.

Since, in the classical case, all limits may be constructed from finite limits (in fact, from equalizers) and products, we therefore need only worry about the $\mathcal{S}$-indexed analogues of infinite products and coproducts. (We shall prove the $\mathcal{S}$-indexed analogue of the theorem just quoted in 2.3.20 below.)

To say that an ordinary category $\mathcal{C}$ has I-indexed products, where I is a set, is to say that the diagonal functor $\mathcal{C} \to \mathcal{C}^I$ has a right adjoint. But to say that $\mathcal{C}$ has all small products implies, more generally, that all the transition functors $x^*: \mathcal{C}^J \to \mathcal{C}^I$ in the naive indexing of $\mathcal{C}$ over **Set** have right adjoints, and that these right adjoints themselves form indexed functors, i.e. they commute with re-indexing in a suitable sense. Because we do not wish to accord any special status to the terminal object of our base category $\mathcal{S}$, we take the latter characterization

as the basis of our definition. Until further notice, we shall assume that our base category $\mathcal{S}$ has pullbacks.

Definition 1.4.2 *(a)* Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category. We say that $\mathbb{C}$ *has $\mathcal{S}$-indexed products* if we are given, for each $x\colon I \to J$ in $\mathcal{S}$, a right adjoint $\Pi_x : \mathcal{C}^I \to \mathcal{C}^J$ for x^*, and for each pullback square

$$\begin{tikzcd} I \arrow[r, "x"] \arrow[d, "y"'] & J \arrow[d, "z"] \\ K \arrow[r, "w"] & L \end{tikzcd}$$

in $\mathcal{S}$, the canonical natural transformation ψ in the diagram

$$\begin{tikzcd} \mathcal{C}^J \arrow[r, "x^*"] \arrow[d, "\Pi_z"'] \arrow[dr, phantom, "\overset{\psi}{\to}"] & \mathcal{C}^I \arrow[d, "\Pi_y"] \\ \mathcal{C}^L \arrow[r, "w^*"] & \mathcal{C}^K \end{tikzcd}$$

(defined by taking $\psi_A : w^*\Pi_z A \to \Pi_y x^* A$ to be the transpose across $(y^* \dashv \Pi_y)$ of the composite $y^*w^*\Pi_z A \cong x^*z^*\Pi_z A \to x^*A$ induced by the counit of $(z^* \dashv \Pi_z)$) is an isomorphism.

(b) We say that $\mathbb{C}$ is *$\mathcal{S}$-complete* if it has $\mathcal{S}$-indexed finite limits and $\mathcal{S}$-indexed products.

The condition on pullback squares is commonly known as the *Beck–Chevalley condition*; cf. A2.2.4. It is often rather sloppily stated as the existence of a natural isomorphism $w^*\Pi_z \cong \Pi_y x^*$, rather than the requirement that the particular natural transformation ψ defined above should be invertible; but in practice, once one has verified the former, it is usually a fair bet that the latter will be true.

Remark 1.4.3 The two conditions '$\mathbb{C}$ has $\mathcal{S}$-indexed products' and '$\mathbb{C}$ has finite limits' of 1.4.2*(b)* are not entirely independent. We shall see in the next section that if the base category $\mathcal{S}$ is positive coherent, and $\mathbb{C}$ is a stack for the coherent coverage on $\mathcal{S}$, then for every $I \in \mathrm{ob}\,\mathcal{S}$ we have $\mathcal{C}^{I \amalg I} \simeq \mathcal{C}^I \times \mathcal{C}^I$, by an equivalence identifying ∇^* (where $\nabla\colon I \amalg I \to I$ is the codiagonal map) with the diagonal functor $\mathcal{C}^I \to \mathcal{C}^I \times \mathcal{C}^I$. So the assertion that the ∇^* have right adjoints is equivalent to saying that the categories $\mathcal{C}^I$ all have binary products; and the

Beck–Chevalley condition for pullback squares of the form

$$\begin{tikzcd} J \amalg J \arrow[r, "\nabla"] \arrow[d, "x \amalg x"'] & J \arrow[d, "x"] \\ I \amalg I \arrow[r, "\nabla"] & I \end{tikzcd}$$

is equivalent to saying that these binary products are preserved by the transition functors x^*. Similar arguments involving the initial object of $\mathcal{S}$ ensure that, if $\mathbb{C}$ has $\mathcal{S}$-indexed products, then its fibres have terminal objects which are preserved by the transition functors. Thus, under the above hypotheses on $\mathcal{S}$ and $\mathbb{C}$, the condition '$\mathbb{C}$ has finite limits' of 1.4.2(b) can be reduced to '$\mathbb{C}$ has equalizers' (or even to '$\mathbb{C}$ has coreflexive equalizers'; cf. A1.2.10).

By the uniqueness of adjoints, it is clear that if the right adjoints Π_x exist for all x they are (covariantly) pseudo-functorial in x. At first sight, this gives another way of defining a natural transformation in the square appearing in 1.4.2(a), as the transpose across $(w^* \dashv \Pi_w)$ of the composite $\Pi_z A \to \Pi_z \Pi_x x^* A \cong \Pi_w \Pi_y x^* A$ induced by the unit of $(x^* \dashv \Pi_x)$, but in fact this transformation is equal to ψ – this follows from the way in which the coherence isomorphism $\Pi_z \Pi_x \cong \Pi_w \Pi_y$ is constructed from $y^* w^* \cong x^* z^*$.

Of course, we say that $\mathbb{C}$ *has $\mathcal{S}$-indexed coproducts* if each x^* has a left adjoint $\Sigma_x \colon C^I \to C^J$, and the Beck–Chevalley condition holds for pullbacks in $\mathcal{S}$.

Examples 1.4.4 (a) To show that the Beck–Chevalley condition is not always a triviality, consider the **Set**-indexed category $\mathbb{T}$ of 1.2.2(h). We have already observed that its transition functors have left adjoints; but in the pullback square

$$\begin{tikzcd} 0 \arrow[r, "x"] \arrow[d, "y"'] & 1 \arrow[d, "z"] \\ 2 \arrow[r, "w"] & 2 \end{tikzcd}$$

(where n denotes an n-element set, and z and w are constant maps with disjoint images), the composite $w^* z_*$ sends the unique topology on 1 to the indiscrete topology on 2, whereas $y_* x^*$ sends it to the discrete topology.

(b) A rather more substantial example of the same phenomenon is given by the **Cat**-indexed category which to a small category $\mathcal{C}$ associates the functor category $[\mathcal{C}, \mathbf{Set}]$, and to a functor $f \colon \mathcal{C} \to \mathcal{D}$ the functor $f^* \colon [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ obtained by composition with f. As we observed in A4.1.4, these functors have adjoints on both sides; but they fail to satisfy the Beck–Chevalley condition in

general. For a simple counterexample, consider any pullback square of the form

$$\begin{tikzcd} \mathcal{C} \arrow[r, "1"] \arrow[d, "1"'] & \mathcal{C} \arrow[d, "i"] \\ \mathcal{C} \arrow[r, "i"] & \mathcal{D} \end{tikzcd}$$

where $\mathcal{C}$ is a non-full subcategory of $\mathcal{D}$: the composite (for either left or right adjoints) around the top and left sides of the square is clearly the identity functor, but the composite $i^* \circ \varinjlim_i$ is the identity only if i is full and faithful (cf. A4.2.12(b)). Another counterexample is provided by the pullback

$$\begin{tikzcd} \emptyset \arrow[r] \arrow[d] & \mathbf{1} \arrow[d, "1"] \\ \mathbf{1} \arrow[r, "0"] & \mathbf{2} \end{tikzcd}$$

where $\mathbf{n}$ denotes an n-element totally ordered set and $0, 1$ are the bottom and top elements of $\mathbf{2}$; it is easily checked that $1^* \circ \varinjlim_0$ is the identity functor $\mathbf{Set} \to \mathbf{Set}$, but going the other way round the Beck–Chevalley square produces the constant functor with value 0.

The next result generalizes a phenomenon which we observed in 1.3.7(a) and (d).

Lemma 1.4.5 *Let $\Pi\colon \mathcal{C} \to \mathcal{S}$ be a fibration, and $\mathbb{C}$ the corresponding $\mathcal{S}$-indexed category. Then*

(i) *the transition functors of $\mathbb{C}$ have left adjoints iff Π is also an opfibration;*

(ii) *$\mathbb{C}$ has $\mathcal{S}$-indexed coproducts iff Π is an opfibration and the pullbacks of supine morphisms of $\mathcal{C}$ along prone morphisms (exist and) are supine.*

Proof (i) Suppose the x^* have left adjoints Σ_x. Then, for each $x\colon I \to J$ in $\mathcal{S}$ and each $A \in \text{ob } \mathcal{C}^I$, we have a morphism $\tau(A, x)\colon A \to \Sigma_x A$ in $\mathcal{C}$ obtained by composing the unit map $A \to x^*\Sigma_x A$ in $\mathcal{C}^I$ with the prone morphism $\sigma(\Sigma_x A, x)$. Now, if we are given $f\colon A \to B$ in $\mathcal{C}$ such that Πf is the composite of x with $y\colon J \to K$, then we obtain a morphism $A \to (yx)^*B \cong x^*y^*B$ in $\mathcal{C}^I$ and hence a morphism $\Sigma_x A \to y^*B$ in $\mathcal{C}^J$; composing this with $\sigma(B, y)$ yields the unique factorization of f through $\tau(A, x)$ lying over y. Thus $\tau(A, x)$ is supine, and so Π is an opfibration. Conversely, if Π is an opfibration, then by defining $\Sigma_x A$ to be the codomain of a supine lifting of x with domain A, we obtain an adjunction $(\Sigma_x \dashv x^*)$, since morphisms $\Sigma_x A \to B$ in $\mathcal{C}^J$ and morphisms $A \to x^*B$ in $\mathcal{C}^I$ both correspond bijectively to morphisms $A \to B$ in $\mathcal{C}$ lying over x.

(ii) Given (i), we have to show that the Beck–Chevalley condition for the Σ_x is equivalent to the stability of supine morphisms under pullback along prone morphisms. Given a pullback square

$$\begin{tikzcd} I \arrow[r, "x"] \arrow[d, "y"] & J \arrow[d, "z"] \\ K \arrow[r, "w"] & L \end{tikzcd}$$

in $\mathcal{S}$ and an object A of $\mathcal{C}^J$, we may form a commutative square

$$\begin{tikzcd} x^*A \arrow[r, "\sigma(A,x)"] \arrow[d, "h"] & A \arrow[d, "\tau(A,z)"] \\ w^*\Sigma_z A \arrow[r, "\sigma(\Sigma_z A, w)"] & \Sigma_z A \end{tikzcd}$$

in $\mathcal{C}$, where h is the unique factorization of the composite $\tau(A,z)\sigma(A,x)$ through $\sigma(\Sigma_z A, w)$ lying over y. By an argument like that of 1.3.3(ii), it is readily seen that this square is a pullback in $\mathcal{C}$. So the assertion that its left edge h is supine is equivalent to saying that we have a canonical isomorphism $\Sigma_y x^* A \cong w^*\Sigma_z A$ – and this isomorphism is necessarily the natural map appearing in the Beck–Chevalley condition. □

The following lemma is a useful labour-saving device:

Lemma 1.4.6 *Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category, and suppose that, for each $x\colon I \to J$ in $\mathcal{S}$, $x^*\colon \mathcal{C}^J \to \mathcal{C}^I$ has both left and right adjoints Σ_x and Π_x. Then $\mathbb{C}$ has $\mathcal{S}$-indexed products iff it has $\mathcal{S}$-indexed coproducts.*

Proof Suppose that $\mathbb{C}$ has $\mathcal{S}$-indexed products, i.e. that the Beck–Chevalley condition holds for the right adjoints. Then, if

$$\begin{tikzcd} I \arrow[r, "x"] \arrow[d, "y"] & J \arrow[d, "z"] \\ K \arrow[r, "w"] & L \end{tikzcd}$$

is a pullback square in $\mathcal{S}$, the functors in the diagram

$$\begin{tikzcd} \mathcal{C}^J \arrow[r, "x^*"] \arrow[d, "\Sigma_z"] & \mathcal{C}^I \arrow[d, "\Sigma_y"] \\ \mathcal{C}^L \arrow[r, "w^*"] & \mathcal{C}^K \end{tikzcd}$$

all have right adjoints, and the latter commute up to isomorphism; so the square above commutes up to isomorphism. Moreover, the canonical transformation $\Sigma_y x^* \to w^*\Sigma_z$ can be shown to be the 'mate' under the adjunctions of the canonical transformation $z^*\Pi_w \to \Pi_x y^*$, so the first is an isomorphism iff the second is. □

The alert reader will recognize that we have seen a particular case of 1.4.6 before, in A1.4.11: there we observed that the indexed category $\mathbb{S}\mathrm{ub}_{\mathcal{S}}$ of 1.2.2(g) has $\mathcal{S}$-indexed coproducts iff $\mathcal{S}$ is a regular category (cf. A1.3.1(iii)), and deduced that it has $\mathcal{S}$-indexed products provided its transition functors have right adjoints, i.e. provided $\mathcal{S}$ is a Heyting category. Note also that $\mathbb{S}\mathrm{ub}_{\mathcal{S}}$ is $\mathcal{S}$-cocomplete iff $\mathcal{S}$ is a coherent category: since its fibres are preorders, the only finite colimits we have to worry about are coproducts (i.e. unions).

Similarly, we have

Lemma 1.4.7 *Let $\mathcal{S}$ be a cartesian category, and let $\mathbb{S}$ be the canonical indexing of $\mathcal{S}$ over itself, as defined in* 1.2.2(c). *Then*

(i) $\mathbb{S}$ *always has $\mathcal{S}$-indexed coproducts;*

(ii) $\mathbb{S}$ *is $\mathcal{S}$-cocomplete iff $\mathcal{S}$ has finite colimits which are stable under pullback;*

(iii) $\mathbb{S}$ *has $\mathcal{S}$-indexed products iff it is $\mathcal{S}$-complete, iff $\mathcal{S}$ is locally cartesian closed.*

Proof (i) We know that the pullback functor $x^* : \mathcal{S}/J \to \mathcal{S}/I$ has a left adjoint Σ_x, by A1.2.8; and the Beck–Chevalley condition for the square

$$\begin{tikzcd} \mathcal{S}/J \arrow[r, "x^*"] \arrow[d, "\Sigma_z"] & \mathcal{S}/I \arrow[d, "\Sigma_y"] \\ \mathcal{S}/L \arrow[r, "w^*"] & \mathcal{S}/K \end{tikzcd}$$

is simply the assertion that a composite of two pullback squares is a pullback.

(ii) is immediate from (i) and the fact that the forgetful functors $\mathcal{S}/I \to \mathcal{S}$ preserve and reflect colimits.

(iii) A1.5.3 tells us that $\mathcal{S}$ is locally cartesian closed iff the x^* all have right adjoints Π_x, and the Beck–Chevalley condition for the Π_x follows from that for the Σ_x, by 1.4.6. Finally, since $\mathcal{S}$ is cartesian, the slice categories $\mathcal{S}/I$ all have finite limits, and the x^* preserve them since they have left adjoints. □

In particular, we note that $\mathbb{S}$ is both $\mathcal{S}$-complete and $\mathcal{S}$-cocomplete iff $\mathcal{S}$ is locally cartesian closed and cocartesian.

The next result is a useful extension of 1.4.7.

Lemma 1.4.8 *Let $F: \mathcal{S} \to \mathcal{C}$ be a functor between cartesian categories, and let $\mathbb{C}$ denote the $\mathcal{S}$-indexed category $I \mapsto \mathcal{C}/F(I)$ of* 1.2.2(*d*). *The following are equivalent:*

(i) *F preserves pullbacks;*

(ii) *$\mathbb{C}$ has $\mathcal{S}$-indexed coproducts;*

(iii) *(if $\mathcal{C}$ is locally cartesian closed) $\mathbb{C}$ has $\mathcal{S}$-indexed products.*

Proof As before, the pullback functors $F(x)^*: \mathcal{C}/F(J) \to \mathcal{C}/F(I)$ automatically have left adjoints, and they have right adjoints if $\mathcal{C}$ is locally cartesian closed; the question is thus whether the Beck–Chevalley condition holds. If F preserves pullbacks, then the condition holds for the same reason as in 1.4.7. Conversely, if the condition holds for a pullback square

$$\begin{tikzcd} I \arrow[r, "x"] \arrow[d, "y"] & J \arrow[d, "z"] \\ K \arrow[r, "w"] & L \end{tikzcd}$$

in $\mathcal{S}$, then on chasing the terminal object $1_{F(J)}$ of $\mathcal{C}/F(J)$ around the square

$$\begin{tikzcd} \mathcal{C}/F(J) \arrow[r, "x^*"] \arrow[d, "\Sigma_z"] & \mathcal{C}/F(I) \arrow[d, "\Sigma_y"] \\ \mathcal{C}/F(L) \arrow[r, "w^*"] & \mathcal{C}/F(K) \end{tikzcd}$$

we see that F must preserve this pullback. □

We recall from A1.2.9 that any pullback-preserving functor $F\colon \mathcal{S} \to \mathcal{C}$ can be factored as

$$\mathcal{S} \xrightarrow{\hat{F}} \mathcal{C}/F1 \longrightarrow \mathcal{C}$$

where $\hat{F}$ is cartesian; and the $\mathcal{S}$-indexed category obtained from $\hat{F}$ via 1.2.2(d) coincides with that obtained from F. So there is no real loss of generality, in this context, in assuming that F is cartesian.

Lemma 1.4.8 can be further generalized:

Lemma 1.4.9 *Let $F\colon \mathcal{T} \to \mathcal{S}$ be a functor between cartesian categories. Then the change of base functor $F^*\colon \mathfrak{CAT}_{\mathcal{S}} \to \mathfrak{CAT}_{\mathcal{T}}$ of* 1.2.2(e) *preserves the property of having indexed coproducts (and products) iff F preserves pullbacks.*

Proof If $\mathbb{C}$ is a $\mathcal{T}$-indexed category with coproducts, then for any $x\colon I \to J$ in $\mathcal{S}$ the transition functor $x^*\colon (F^*\mathcal{C})^J \to (F^*\mathcal{C})^I$ coincides with $F(x)^*\colon \mathcal{C}^{FJ} \to \mathcal{C}^{FI}$, and so has a left adjoint; and if F preserves pullbacks then the Beck–Chevalley squares for $F^*\mathbb{C}$ are particular cases of those for $\mathbb{C}$, and so they commute. The argument for products is similar.

Conversely, suppose F^* preserves the property of having indexed coproducts. Then in particular $F^*\mathbb{S}$ has $\mathcal{T}$-indexed coproducts by 1.4.7, where $\mathbb{S}$ is the canonical indexing of $\mathcal{S}$ over itself; but this is the indexing of $\mathcal{S}$ over $\mathcal{T}$ obtained from F as in 1.2.2(d), so by 1.4.8 F must preserve pullbacks. □

Of course, any change-of-base functor preserves the property of having (indexed) finite limits or colimits; so we deduce that if F preserves pullbacks then F^* preserves indexed completeness and cocompleteness. In particular, since functors of the form $(-) \times I\colon \mathcal{S} \to \mathcal{S}$ preserve pullbacks, we note that an $\mathcal{S}$-indexed category of the form $\mathbb{C}^I$ (as defined in 1.2.2(f)) is $\mathcal{S}$-complete or cocomplete if $\mathbb{C}$ is.

It is possible to characterize the $\mathcal{S}$-indexed categories arising from cartesian functors $F\colon \mathcal{S} \to \mathcal{C}$ in fibrational terms: the following argument is due to J. L. Moens [820]. First, we need a couple of new definitions. We say that an $\mathcal{S}$-indexed category $\mathbb{C}$ has *stable $\mathcal{S}$-indexed coproducts* if the supine morphisms in the corresponding fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$ are stable under pullback along arbitrary morphisms, not just along prone ones (cf. 1.4.5).

Lemma 1.4.10 *Let $\Pi\colon \mathcal{C} \to \mathcal{S}$ be a cartesian fibration, and $\mathbb{C}$ the corresponding $\mathcal{S}$-indexed category. Then $\mathbb{C}$ has stable $\mathcal{S}$-indexed coproducts iff the following condition is satisfied: suppose given $x\colon I \to J$ in $\mathcal{S}$, an object A of $\mathcal{C}^I$ and a morphism $B \to \Sigma_x A$ in $\mathcal{C}^J$; then if we form the pullback*

$$\begin{array}{ccc} C & \longrightarrow & x^*B \\ \downarrow & & \downarrow \\ A & \longrightarrow & x^*\Sigma_x A \end{array}$$

in C^I (where the bottom edge of the square is the unit of $(\Sigma_x \dashv x^*)$), the transpose $\Sigma_x C \to B$ of the top edge is an isomorphism.

Proof By 1.4.5(ii), if $\mathbb{C}$ has $\mathcal{S}$-indexed coproducts, they are stable iff the supine morphisms in the total category of the corresponding fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$ are stable under pullback along vertical morphisms. But in this category we may form the diagram

$$\begin{tikzcd} C \arrow[r] \arrow[d] & x^*B \arrow[r, "\sigma(B,x)"] \arrow[d] & B \arrow[d] \\ A \arrow[r] & x^*\Sigma_x A \arrow[r, "\sigma(\Sigma_x A, x)"] & \Sigma_x A \end{tikzcd}$$

where the right-hand square is a pullback by 1.3.3(ii), and the bottom composite is the supine morphism $\tau(A, x)$. So the outer rectangle is a pullback iff the left-hand square is; and the top composite is supine iff the transpose of $C \to x^*B$ is an isomorphism. □

Given this result, it is easy to see that the naive indexing of an ordinary cartesian category over **Set** has stable **Set**-indexed coproducts iff the ordinary category has coproducts which are stable under pullback, in the usual sense.

We may also define disjointness of coproducts for indexed categories: given a cartesian fibration $\Pi\colon \mathcal{C} \to \mathcal{S}$, we say the corresponding $\mathcal{S}$-indexed category $\mathbb{C}$ *has disjoint $\mathcal{S}$-indexed coproducts* if (it has $\mathcal{S}$-indexed coproducts and) for any supine morphism $f\colon A \to B$ in $\mathcal{C}$, the diagonal morphism $A \to A \times_B A$ is also supine. For the naive indexing of an ordinary cartesian category over **Set**, it is again not hard to see that this is equivalent to disjointness of coproducts as we defined it in A1.4.4.

Lemma 1.4.11 *Let $\Pi\colon \mathcal{C} \to \mathcal{S}$ be a cartesian fibration, and suppose the corresponding $\mathcal{S}$-indexed category has disjoint and stable $\mathcal{S}$-indexed coproducts. Then, for any $x\colon I \to J$ in $\mathcal{S}$ and any $f\colon A \to B$ in C^I, the canonical square*

$$\begin{tikzcd} A \arrow[r] \arrow[d, "f"'] & \Sigma_x(A) \arrow[d, "\Sigma_x(f)"] \\ B \arrow[r] & \Sigma_x(B) \end{tikzcd}$$

is a pullback in $\mathcal{C}$.

Proof It is easy to see that we have a pullback square

$$\begin{tikzcd} A \arrow[r, "{(f,1)}"] \arrow[d, "f"'] & B \times_{\Sigma_x B} A \arrow[d, "1 \times f"] \\ B \arrow[r, "\Delta"] & B \times_{\Sigma_x B} B \end{tikzcd}$$

and hence that the top edge of this square is supine. But we also have a supine morphism $B \times_{\Sigma_x B} A \to B \times_{\Sigma_x B} \Sigma_x A$ lying over the first projection $I \times_J I \to I$ (since this is a pullback of $A \to \Sigma_x A$), and so the composite is a supine morphism $A \to B \times_{\Sigma_x B} \Sigma_x A$ lying over the identity morphism on I. But any such morphism must be an isomorphism. □

Theorem 1.4.12 *Let $\Pi: C \to S$ be a cartesian fibration, and let $\mathbb{C}$ be the corresponding S-indexed category. Then $\mathbb{C}$ has disjoint and stable S-indexed coproducts iff there exists a cartesian functor $F: S \to T$ such that $\mathbb{C}$ is equivalent to the S-indexed category induced by F as in* 1.2.2(d).

Proof Suppose there is such an F. Then we may identify C with the glueing category $\mathbf{Gl}(F)$, as we observed in 1.3.7(b); and it is not hard to see that a morphism

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d] & B \arrow[d] \\ FI \arrow[r, "Fx"] & FJ \end{tikzcd}$$

in this category is supine iff f is an isomorphism. Given this, it is easy to verify the conditions for disjointness and stability.

Conversely, suppose $\mathbb{C}$ has disjoint and stable coproducts. We take T to be the fibre C^1, and we define $F: S \to T$ on objects by $FI = \Sigma_I(1_I)$, where 1_I denotes the terminal object of C^I. For a morphism $x: I \to J$ of S, Fx corresponds to the composite

$$1_I \cong x^*(1_J) \xrightarrow{x^*(\eta)} x^* J^* \Sigma_J(1_J) \cong I^* \Sigma_J(1_J),$$

where η is the unit of $(\Sigma_J \dashv J^*)$; equivalently, it is the unique vertical morphism $FI \to FJ$ making

$$\begin{tikzcd} 1_I \arrow[r] \arrow[d, "{\sigma(1_J, x)}"'] & FI \arrow[d, "Fx"] \\ 1_J \arrow[r] & FJ \end{tikzcd}$$

commute. It is easy to verify that this is functorial. We have a functor $\Phi^I : \mathcal{C}^I \to \mathcal{T}/FI$ which is simply Σ_I applied to morphisms with codomain 1_I, and we have a functor $\Psi^I : \mathcal{T}/FI \to \mathcal{C}^I$ which sends $(f : A \to FI = \Sigma_I(1_I))$ to the pullback

$$
\begin{tikzcd}
\Psi^I(f) \arrow[r] \arrow[d] & A \arrow[d] \\
1_I \arrow[r] & FI
\end{tikzcd}
$$

in $\mathcal{C}$. To verify that Φ is an $\mathcal{S}$-indexed functor, consider the cube

$$
\begin{tikzcd}
x^*A \arrow[rr] \arrow[dr, "\sigma(A,x)"] \arrow[dd] & & (Fx)^*\Sigma_I(A) \arrow[dr] \arrow[dd] & \\
& A \arrow[rr] \arrow[dd] & & \Sigma_I(A) \arrow[dd, "\Phi^I(A)"] \\
1_J \arrow[rr] \arrow[dr, "\sigma(1_I,x)"'] & & FJ \arrow[dr, "Fx"'] & \\
& 1_I \arrow[rr] & & FI
\end{tikzcd}
$$

induced by a morphism $x : J \to I$ in $\mathcal{S}$. Here the front face is a pullback by 1.4.11, the left face is a pullback by 1.3.3(ii), and the right face is a pullback by definition. So the back face is a pullback, and hence its top edge is supine; thus we have $\Phi^J x^*(A) \cong x^*\Phi^I(A)$. The fact that Ψ is an $\mathcal{S}$-indexed functor is an easy application of 1.3.3(ii).

Since the top edge of the pullback square defining $\Psi^I(f)$ is supine, it is clear that the composite $\Phi^I\Psi^I$ is naturally isomorphic to the identity. The fact that $\Psi^I\Phi^I$ is isomorphic to the identity follows from 1.4.11.

Finally, since we now know that the $\mathcal{S}$-indexed category $I \mapsto \mathcal{T}/FI$ is equivalent to $\mathbb{C}$, it has $\mathcal{S}$-indexed coproducts; so F preserves pullbacks by 1.4.8. But it is clear from the construction that it also preserves the terminal object; so it is a cartesian functor. $\square$

Next, we introduce the notion of continuity (= preservation of limits) for $\mathcal{S}$-indexed functors.

Definition 1.4.13 If $F: \mathbb{C} \to \mathbb{D}$ is an indexed functor between complete $\mathcal{S}$-indexed categories, we say F is *$\mathcal{S}$-continuous* (or simply continuous) if each F^I preserves finite limits, and for each $x: I \to J$ in $\mathcal{S}$ the canonical natural transformation $F^J \Pi_x \to \Pi_x F^I$, which is the 'mate' of the isomorphism $x^* F^J \cong F^I x^*$, is an isomorphism.

Lemma 1.4.14 *Let $\mathbb{C}$ and $\mathbb{D}$ be complete $\mathcal{S}$-indexed categories, and $F: \mathbb{C} \to \mathbb{D}$ an $\mathcal{S}$-indexed functor having an indexed left adjoint (i.e. such that each F^I has a left adjoint L^I, in such a way that the L^I form an $\mathcal{S}$-indexed functor and the unit and counit of the adjunction are $\mathcal{S}$-indexed natural transformations). Then F is $\mathcal{S}$-continuous.*

Proof Since each F^I has a left adjoint, it preserves the finite limits which exist in $\mathcal{C}^I$. For the second condition, we note that the requirement that the unit and counit of the adjunction be $\mathcal{S}$-indexed forces the isomorphism up to which the diagram

$$\begin{tikzcd} \mathcal{D}^J \arrow[r, "x^*"] \arrow[d, "L^J"] & \mathcal{D}^I \arrow[d, "L^I"] \\ \mathcal{C}^J \arrow[r, "x^*"] & \mathcal{C}^I \end{tikzcd}$$

commutes to be the mate of that up to which the corresponding diagram for F commutes. So, on taking the right adjoints of all the functors in this diagram, we obtain a diagram

$$\begin{tikzcd} \mathcal{C}^I \arrow[r, "\Pi_x"] \arrow[d, "F^I"] & \mathcal{C}^J \arrow[d, "F^J"] \\ \mathcal{D}^I \arrow[r, "\Pi_x"] & \mathcal{D}^J \end{tikzcd}$$

commuting up to an isomorphism, which again must be the mate (across a different pair of adjunctions!) of the given isomorphism $x^* F^J \cong F^I x^*$. □

Under suitable 'smallness' conditions on the categories $\mathbb{C}$ and $\mathbb{D}$, Lemma 1.4.14 has a converse (the Indexed Adjoint Functor Theorem), which we shall prove in 2.4.6 below.

The notion of indexed monad on an indexed category is self-explanatory; it is also easy to see that an indexed monad $\mathbb{T} = (T, \eta, \mu)$ on an $\mathcal{S}$-indexed category $\mathbb{C}$ gives rise to an indexed category of algebras $\mathbb{C}^{\mathbb{T}}$, whose fibre over I is the category of algebras for the monad (T^I, η^I, μ^I) on $\mathcal{C}^I$. The following result, though unsurprising, will be of importance later on.

Lemma 1.4.15 *Let $\mathbb{T}$ be an indexed monad on an $\mathcal{S}$-indexed category $\mathbb{C}$.*

(i) *If $\mathbb{C}$ has $\mathcal{S}$-indexed products, so has $\mathbb{C}^{\mathbb{T}}$, and the indexed forgetful functor $\mathbb{C}^{\mathbb{T}} \to \mathbb{C}$ preserves them.*

(ii) *If $\mathbb{C}$ has $\mathcal{S}$-indexed coproducts and T preserves them, then $\mathbb{C}^{\mathbb{T}}$ has $\mathcal{S}$-indexed coproducts and the forgetful functor preserves them.*

Proof (i) Let $x: I \to J$ be a morphism of $\mathcal{S}$, and (A, α) a $\mathbb{T}^I$-algebra in $\mathcal{C}^I$. Consider the composite

$$T^J \Pi_x(A) \xrightarrow{\theta_A} \Pi_x T^I(A) \xrightarrow{\Pi_x(\alpha)} \Pi_x(A),$$

where θ is the mate of the natural isomorphism $x^* T^J \cong T^I x^*$. It is straightforward to verify that this composite is a $\mathbb{T}^J$-algebra structure on $\Pi_x(A)$; and that, given any $\mathbb{T}^J$-algebra (B, β), a morphism $B \to \Pi_x(A)$ in $\mathcal{C}^J$ is a $\mathbb{T}^J$-algebra homomorphism iff its transpose $x^*(B) \to A$ is a $\mathbb{T}^I$-algebra homomorphism. Thus we have constructed a right adjoint for $x^*: (\mathcal{C}^J)^{\mathbb{T}^J} \to (\mathcal{C}^I)^{\mathbb{T}^I}$. The fact that the forgetful functors from $\mathbb{T}$-algebras to objects commute with these right adjoints is immediate from the definitions. Finally, the Beck–Chevalley condition for $\mathbb{C}^{\mathbb{T}}$ can be 'lifted' from that for $\mathbb{C}$, since the forgetful functors reflect isomorphisms.

(ii) Given a morphism $x: I \to J$ and a $\mathbb{T}^I$-algebra (A, α) in $\mathcal{C}^I$, we consider the morphism

$$T^J \Sigma_x(A) \cong \Sigma_x T^I(A) \xrightarrow{\Sigma_x(\alpha)} \Sigma_x(A)\,.$$

It is again straightforward to verify that this is a $\mathbb{T}^J$-algebra structure on $\Sigma_x(A)$, and that this construction yields a functor $(\mathcal{C}^I)^{\mathbb{T}^I} \to (\mathcal{C}^J)^{\mathbb{T}^J}$ left adjoint to x^*. Once again, the Beck–Chevalley condition for $\mathbb{C}^{\mathbb{T}}$ may be lifted from that for $\mathbb{C}$. □

Given an arbitrary $\mathcal{S}$-indexed category $\mathbb{C}$, we may freely adjoin $\mathcal{S}$-indexed coproducts to it, by the following construction. Let $\mathrm{Fam}(\mathbb{C})$, the category of ($\mathcal{S}$-indexed) *families* of objects of $\mathbb{C}$, be the $\mathcal{S}$-indexed category whose I-indexed families of objects are pairs (x, A) where $x: J \to I$ is a morphism with codomain I in $\mathcal{S}$ and A is an object of $\mathcal{C}^J$, and whose morphisms $(x, A) \to (x', A')$ are pairs (y, f) where $y: x \to x'$ in $\mathcal{C}/I$ and $f: A \to y^* A'$ in $\mathcal{C}^J$. (Note in particular that $\mathrm{Fam}(\mathcal{C})^1$ is simply the total category of the fibration corresponding to $\mathbb{C}$.) If $z: K \to I$ is a morphism of $\mathcal{C}$, then the re-indexing functor $z^*: \mathrm{Fam}(\mathcal{C})^I \to \mathrm{Fam}(\mathcal{C})^K$ sends (x, A) to $(\pi_1, \pi_2^* A)$ where the π_i are the projections from the pullback $K \times_I J$. The left adjoint Σ_z of z^* is given by composition: $\Sigma_z(w, B) = (zw, B)$. It is straightforward to verify that the above formulae do define an $\mathcal{S}$-indexed category with $\mathcal{S}$-indexed coproducts (the verification of the Beck–Chevalley condition is like that in the proof of 1.4.7(i) – indeed, 1.4.7(i) is a special case of it, since the canonical indexing $\mathbb{S}$ is none other than $\mathrm{Fam}(\mathbf{1})$,

where $\mathbf{1}$ is the indexed category with one object and one morphism in each fibre); and we have

Proposition 1.4.16 *Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category.*

(i) *There is an $\mathcal{S}$-indexed full embedding $\mathbb{C} \to \mathrm{Fam}(\mathbb{C})$.*

(ii) *If $\mathbb{D}$ is an $\mathcal{S}$-indexed category with $\mathcal{S}$-indexed coproducts, then any $\mathcal{S}$-indexed functor $F\colon \mathbb{C} \to \mathbb{D}$ extends, uniquely up to canonical isomorphism, to a coproduct-preserving functor $\tilde{F}\colon \mathrm{Fam}(\mathbb{C}) \to \mathbb{D}$.*

(iii) *The assignment $\mathbb{C} \mapsto \mathrm{Fam}(\mathbb{C})$ has the structure of a KZ-monad (cf. 1.1.11) on $\mathfrak{Cat}_{\mathcal{S}}$; and the algebras for this monad are exactly the $\mathcal{S}$-indexed categories with $\mathcal{S}$-indexed coproducts.*

Proof (i) The embedding is defined on each fibre C^I by $A \mapsto (1_I, A)$; it is straightforward to verify that this works.

(ii) Again, it is straightforward to verify that $\tilde{F}$ may be defined (and, up to isomorphism, must be defined) by $\tilde{F}^I(x, A) = \Sigma_x(F^J(A))$.

(iii) In particular, if $\mathbb{D}$ has $\mathcal{S}$-indexed coproducts, then the identity functor on $\mathbb{D}$ must extend to a functor $\mathrm{Fam}(\mathbb{D}) \to \mathbb{D}$, sending (x, A) to $\Sigma_x(A)$. It is easy to verify that this functor is an indexed left adjoint for the embedding of (i): the unit of the adjunction at an object (x, A) is given by $(x, \eta_A)\colon (x, A) \to (1_I, \Sigma_x(A))$ where η is the unit of $(\Sigma_x \dashv x^*)$. Applying this to the case $\mathbb{D} = \mathrm{Fam}(\mathbb{C})$, we obtain the multiplication of the monad, together with the fact that it is a KZ-monad. Then by 1.1.13 we know that the Fam-algebras are the indexed categories for which the embedding of (i) has a left adjoint; but these are exactly the categories with $\mathcal{S}$-indexed coproducts. □

Remark 1.4.17 If our base category $\mathcal{S}$ is a topos with a natural number object, and we merely wish to adjoin finite coproducts to an $\mathcal{S}$-indexed category $\mathbb{C}$, then we may do so by forming the full subcategory $\mathrm{Fam}_f(\mathbb{C})$ of $\mathrm{Fam}(\mathbb{C})$ whose objects are pairs (x, A) such that $x\colon J \to I$ is a finite cardinal in $\mathcal{S}/I$ (cf. A2.5.14). (More generally, given any particular morphism $x\colon X \to I$ in $\mathcal{S}$, we could form the indexed category obtained by 'freely adjoining coproducts indexed by fibres of x'; the particular case under discussion is that in which x is the generic finite cardinal in $\mathcal{S}/N$.) By a result to be proved in D5.2.12, if we are given a composable pair of morphisms $x\colon J \to I$, $y\colon K \to J$ in $\mathcal{S}$ such that x is a finite cardinal in $\mathcal{S}/I$ and y is a finite cardinal in $\mathcal{S}/J$, then the composite xy is (isomorphic to) a finite cardinal in $\mathcal{S}/I$; hence $\mathrm{Fam}_f(\mathbb{C})$ does indeed have coproducts indexed by finite cardinals in $\mathcal{S}$ (that is, if $x\colon J \to I$ is a finite cardinal in $\mathcal{S}/I$, then $x^*\colon \mathrm{Fam}_f(C)^I \to \mathrm{Fam}_f(C)^J$ has a left adjoint, and the Beck–Chevalley condition holds for all pullback squares for which it makes sense). Moreover, the proof of 1.4.16 can easily be adapted to show that $\mathrm{Fam}_f(\mathbb{C})$ is the free category with finite-cardinal-indexed coproducts generated by $\mathbb{C}$.

Of course, we could also regard the problem of adjoining finite coproducts to an indexed category as one to be tackled 'fibrewise', using the 'external' analogue

of the $\mathrm{Fam}_f(-)$ construction on each fibre. By the argument of 1.4.3 above, this will produce the same result (up to equivalence) if the category $\mathbb{C}$ is a stack for the coherent coverage on $\mathcal{S}$. (See also D5.2.13.)

Suggestions for further reading: Moens [820], Street [1137], Wood [1232].

B1.5 Descent conditions and stacks

One problem with the notion of indexed category is that it is a little too 'relaxed', in the sense that the relationships between the fibres C^I for different I are not strictly enough controlled. We have already observed the effect of this in 1.2.2(b) above: if we take our base category $\mathcal{S}$ to be **Set**, then since 1 is a generator for **Set** we might expect the entire structure of a **Set**-indexed category $\mathbb{C}$ to be controlled (up to equivalence) by the fibre C^1, but we have seen that a given ordinary category may admit indexings over **Set** which are genuinely different from the naive one.

One way of excluding at least some of the 'pathological' possibilities for indexings is to require them to satisfy descent conditions for some given coverage on $\mathcal{S}$. (Here we are using the term 'coverage' in the sense in which it was defined in A2.1.9 (or C2.1.1); however, since we shall be concerned exclusively with base categories having finite limits (and in particular pullbacks), we shall generally assume that our coverages satisfy the strong form (C') of the 'pullback-stability' condition rather than the original weak form (C).) The idea of a descent condition is that, given a covering family $(x_\gamma : J_\gamma \to I \mid \gamma \in \Gamma)$ in $\mathcal{S}$, we should be able to 'reconstruct' the fibre C^I up to equivalence from the categories C^{J_γ}, together with additional 'descent data'.

To explain the form which the descent data takes, consider an object A of C^I, and let A_γ denote the object $x_\gamma^*(A)$ of C^{J_γ}, for each $\gamma \in \Gamma$. For each pair (γ, δ) of (not necessarily distinct) indices in Γ, we may form the pullback

$$\begin{tikzcd} K_{\gamma\delta} \arrow[r, "y_{\gamma\delta}"] \arrow[d, "z_{\gamma\delta}"] & J_\gamma \arrow[d, "x_\gamma"] \\ J_\delta \arrow[r, "x_\delta"] & I \end{tikzcd}$$

in $\mathcal{S}$; then since the above square commutes we have a canonical isomorphism $f_{\gamma\delta} : y_{\gamma\delta}^*(A_\gamma) \to z_{\gamma\delta}^*(A_\delta)$ in $C^{K_{\gamma\delta}}$. Moreover, these isomorphisms are compatible, in the following senses:

(a) If $w_\gamma : J_\gamma \to K_{\gamma\gamma}$ is the diagonal map, then $w_\gamma^*(f_{\gamma\gamma})$ is the canonical isomorphism $w_\gamma^* y_{\gamma\gamma}^*(A_\gamma) \cong A_\gamma \cong w_\gamma^* z_{\gamma\gamma}^*(A_\gamma)$ arising from the equations $y_{\gamma\gamma} w_\gamma = 1_{J_\gamma} = z_{\gamma\gamma} w_\gamma$.

(*b*) If we form the 'triple pullback'

$$\begin{tikzcd}
& J_\gamma \arrow[dr, "x_\gamma"] & \\
L_{\gamma\delta\epsilon} \arrow[ur] \arrow[r] \arrow[dr] & J_\delta \arrow[r, "x_\delta"] & I \\
& J_\epsilon \arrow[ur, "x_\epsilon"'] &
\end{tikzcd}$$

together with the three projections $t_{\gamma\delta\epsilon}\colon L_{\gamma\delta\epsilon} \to K_{\gamma\delta}$, $u_{\gamma\delta\epsilon}\colon L_{\gamma\delta\epsilon} \to K_{\gamma\epsilon}$ and $v_{\gamma\delta\epsilon}\colon L_{\gamma\delta\epsilon} \to K_{\delta\epsilon}$, then the diagram

$$\begin{tikzcd}
t^*_{\gamma\delta\epsilon} y^*_{\gamma\delta}(A_\gamma) \arrow[r, "t^*_{\gamma\delta\epsilon}(f_{\gamma\delta})"] \arrow[d] & t^*_{\gamma\delta\epsilon} z^*_{\gamma\delta}(A_\delta) \arrow[r] & v^*_{\gamma\delta\epsilon} y^*_{\delta\epsilon}(A_\delta) \arrow[d, "v^*_{\gamma\delta\epsilon}(f_{\delta\epsilon})"] \\
u^*_{\gamma\delta\epsilon} y^*_{\gamma\epsilon}(A_\gamma) \arrow[r, "u^*_{\gamma\delta\epsilon}(f_{\gamma\epsilon})"] & u^*_{\gamma\delta\epsilon} z^*_{\gamma\epsilon}(A_\epsilon) \arrow[r] & v^*_{\gamma\delta\epsilon} z^*_{\delta\epsilon}(A_\epsilon)
\end{tikzcd}$$

commutes, where the unnamed arrows are the canonical isomorphisms arising from commutative squares in $\mathcal{S}$.

Definition 1.5.1 Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category, and let $R = (x_\gamma\colon J_\gamma \to I \mid \gamma \in \Gamma)$ be a family of morphisms of $\mathcal{S}$ with common codomain. The category $\mathbf{Desc}(\mathbb{C}, R)$ of *descent data for R in $\mathbb{C}$* has as objects pairs $((A_\gamma), (f_{\gamma\delta}))$ where A_γ is an object of $\mathcal{C}^{J_\gamma}$ for each γ and the $f_{\gamma\delta}\colon y^*_{\gamma\delta}(A_\gamma) \to z^*_{\gamma\delta}(A_\delta)$ are morphisms satisfying the compatibility conditions (*a*) and (*b*) above, and as morphisms $((A_\gamma)(f_{\gamma\delta})) \to ((B_\gamma), (g_{\gamma\delta}))$ families $(h_\gamma\colon A_\gamma \to B_\gamma \mid \gamma \in \Gamma)$ such that

$$\begin{tikzcd}
y^*_{\gamma\delta}(A_\gamma) \arrow[r, "f_{\gamma\delta}"] \arrow[d, "y^*_{\gamma\delta}(h_\gamma)"] & z^*_{\gamma\delta}(A_\delta) \arrow[d, "z^*_{\gamma\delta}(h_\delta)"] \\
y^*_{\gamma\delta}(B_\gamma) \arrow[r, "g_{\gamma\delta}"] & z^*_{\gamma\delta}(B_\delta)
\end{tikzcd}$$

commutes for each (γ, δ).

At first sight, Definition 1.5.1 may seem rather arbitrary: it is not clear why the conditions (*a*) and (*b*) should be singled out from all the things we might say

about a family of morphisms $(f_{\gamma\delta})$. However, it can be given a formal justification in terms of 2-categorical (pseudo-)limits: we shall discuss a particular case of this in 3.4.11 below. Note that we do not need to specify, in defining the objects of $\mathbf{Desc}(\mathbb{C}, R)$, that the morphisms $f_{\gamma\delta}$ should be isomorphisms, since this is forced by conditions (a) and (b): if $w_{\gamma\delta}\colon K_{\gamma\delta} \to K_{\delta\gamma}$ is the 'twist' isomorphism in $\mathcal{S}$, then (modulo the appropriate canonical isomorphisms) $w^*_{\gamma\delta}(f_{\delta\gamma})$ is necessarily an inverse for $f_{\gamma\delta}$.

We clearly have a functor $\theta_R\colon C^I \to \mathbf{Desc}(\mathbb{C}, R)$ sending an object A to the family $(x^*_\gamma(A) \mid \gamma \in \Gamma)$ with the structure maps described above, and a morphism $h\colon A \to B$ to the family $(x^*_\gamma(h) \mid \gamma \in \Gamma)$.

Definition 1.5.2 (a) We say an $\mathcal{S}$-indexed category $\mathbb{C}$ *satisfies the descent condition* (resp. *the pre-descent condition*) for a family R of morphisms of $\mathcal{S}$ if the functor θ_R just described is one half of an equivalence of categories $\mathbf{Desc}(\mathbb{C}, R) \simeq C^I$ (resp. is full and faithful).

(b) We say $\mathbb{C}$ is a *stack* (resp. a *pre-stack*) for a coverage T on $\mathcal{S}$ if it satisfies the descent condition (resp. the pre-descent condition) for every covering family in T.

Our use of the terms 'pre-descent' and 'descent' is non-standard: the conventional usage is to call these notions 'descent' and 'effective descent' respectively. However, since the stronger notion is clearly the more important one, it seems appropriate to give it the unadorned name.

A stack may be thought of as a 'higher-dimensional version' of the notion of sheaf for a coverage, introduced in A2.1.9. Indeed, if $\mathbb{C}$ is a discrete indexed category (one such that the fibres C^I are discrete categories), and we ignore questions of size, then $\mathbb{C}$ satisfies the descent condition for a family R iff the functor $(I \mapsto \text{ob}\, C^I)$ satisfies the sheaf axiom for R. However, when $\mathbb{C}$ contains non-identity isomorphisms, this equivalence breaks down: even if the assignment $(I \mapsto \text{ob}\, C^I)$ is strictly functorial (as opposed to pseudofunctorial), the assertion that it is a sheaf does not imply that $\mathbb{C}$ is a stack.

Example 1.5.3 Suppose $\mathcal{S}$ is regular, and let A and B be two objects of $\mathcal{S}$ which are locally but not globally isomorphic: that is, there exists a cover $I \twoheadrightarrow 1$ in $\mathcal{S}$ and an isomorphism $f\colon I^*A \to I^*B$ in $\mathcal{S}/I$, but no isomorphism $A \to B$ in $\mathcal{S}$. (For example, in the topos $[G, \mathbf{Set}]$ where G is a group, we could take A to be G acting on itself by left translations, and B to be G with trivial G-action.) Let $\mathbb{C}$ be the $\mathcal{S}$-indexed category defined by taking C^I to be the monoid of endomorphisms of I^*A; since ob C^I is a singleton for all I, the functor $I \mapsto \text{ob}\, C^I$ is clearly a sheaf for the regular coverage (and indeed for any other coverage on $\mathcal{S}$). If $\mathcal{S}$ is cartesian closed, then $I \mapsto \text{mor}\, C^I$ is also a sheaf, since it is representable by the object A^A. Nevertheless, the composite

$$(I \times I)^*(A) \xrightarrow{\pi_1^*(f)} (I \times I)^*(B) \xrightarrow{\pi_2^*(f^{-1})} (I \times I)^*A$$

satisfies the coherence conditions of 1.5.1 for the family consisting of the single morphism $(I \twoheadrightarrow 1)$, and so defines an object of $\mathbf{Desc}(\mathbb{C}, (I \twoheadrightarrow 1))$. And this object is not isomorphic to the unique object in the image of the functor $\theta_{(I \to 1)} : \mathcal{C}^1 \to \mathbf{Desc}(\mathbb{C}, (I \twoheadrightarrow 1))$, for if it were, then we should have an automorphism g of I^*A such that $\pi_1^*(fg) = \pi_2^*(fg)$, and the composite fg would 'descend' along $I \twoheadrightarrow 1$ to yield an isomorphism $A \cong B$. So $\theta_{(I \to 1)}$ is not an equivalence.

On the other hand, if $\mathbb{C}$ is a pre-stack, then for any pair of objects (A, B) of $\mathcal{C}^1$ the assignment

$$I \mapsto \hom_{\mathcal{C}^I}(I^*A, I^*B)$$

is a sheaf; indeed, the assertion that $\mathbb{C}$ is a pre-stack can be viewed as a 'localized' version (that is, one which applies to all the objects of $\mathbb{C}$ and not just those in the fibre over 1) of the assertion that its hom-sets are sheaves. In particular, if the coverage on $\mathcal{S}$ is subcanonical (that is, such that all the representable functors on $\mathcal{S}$ are sheaves – cf. A2.1.11(a); this condition will be satisfied by all the coverages we wish to consider on $\mathcal{S}$), then any locally small indexed category, as defined in 1.3.12 above, is a pre-stack.

Example 1.5.4 If our base category $\mathcal{S}$ is **Set**, we may take T to be the coverage whose members are all jointly-surjective families of functions. Then it is easy to verify that an indexed category $\mathbb{C}$ is a stack iff it satisfies the descent condition for the family of all maps $1 \to I$, for each set I. In turn, this happens iff the functor

$$\mathcal{C}^I \longrightarrow \prod_{i \in I} \mathcal{C}^1$$

whose ith component is i^* is part of an equivalence of categories, i.e. iff $\mathbb{C}$ is equivalent (as an indexed category) to the naive indexing of $\mathcal{C}^1$.

On a more general base category, there are two particular coverages which will be of interest to us: they are the regular coverage on a regular base category (cf. A2.1.11(a)), and the coherent coverage on a positive coherent base (cf. A2.1.11(b)). We next consider the interpretation of the notion of stack in these two particular cases.

For the regular coverage, each cover consists of a single morphism $x : J \to I$. For such a cover, the notion of descent data has a simpler appearance: an object of $\mathbf{Desc}(\mathbb{C}, (x))$ consists of a single object A of $\mathcal{C}^J$ equipped with a single morphism $f : \pi_1^*(A) \to \pi_2^*(A)$ in $\mathcal{C}^{J \times_I J}$, subject to the 'unit condition' that

$$A \cong \Delta^*\pi_1^*(A) \xrightarrow{\Delta^*(f)} \Delta^*\pi_2^*(A) \cong A$$

is the identity morphism on A, and the 'cocycle condition' that the re-indexings of f along the three projections $J \times_I J \times_I J \to J \times_I J$ fit together (modulo canonical isomorphisms) into a commutative triangle. (As before, these conditions are sufficient to ensure that f is actually an isomorphism, and this is

sometimes explicitly assumed in the definition of descent data.) In a general (cartesian) category $\mathcal{S}$, we say that a morphism x is a *descent morphism* (resp. a *pre-descent morphism*) for a particular $\mathcal{S}$-indexed category $\mathbb{C}$ if $\mathbb{C}$ satisfies the descent condition (resp. the pre-descent condition) for the family whose only member is x. If we simply say 'x is a descent morphism' without mentioning a particular indexed category, we mean that it is a descent morphism for the canonical indexing of $\mathcal{S}$ over itself.

In verifying that a particular morphism is a descent morphism, the following result (due to J. Beck; it was this result that first emphasized the importance of the Beck–Chevalley condition) is often useful.

Proposition 1.5.5 *Suppose $\mathbb{C}$ is an $\mathcal{S}$-indexed category having $\mathcal{S}$-indexed products, and let $x\colon J \to I$ be a morphism of $\mathcal{S}$. Then*

(i) $\mathbf{Desc}(\mathbb{C}, (x))$ *is isomorphic to the category of coalgebras for the comonad $\mathbb{G}$ on $\mathcal{C}^J$ induced by the adjunction $(x^* \dashv \Pi_x)$.*

(ii) *x is a descent morphism for $\mathbb{C}$ iff $x^*\colon \mathcal{C}^I \to \mathcal{C}^J$ is comonadic.*

(iii) *In particular, if $\mathbb{C}$ is $\mathcal{S}$-complete, then x is a descent morphism for $\mathcal{C}$ iff $x^*\colon \mathcal{C}^I \to \mathcal{C}^J$ is conservative.*

Proof (i) If A is an object of $\mathcal{C}^J$, then morphisms $f\colon \pi_1^*(A) \to \pi_2^*(A)$ in $\mathcal{C}^{J\times_I J}$ correspond bijectively to morphisms $A \to \Pi_{\pi_1}\pi_2^*(A)$ in $\mathcal{C}^J$; but by the Beck–Chevalley condition the latter correspond bijectively to morphisms $\overline{f}\colon A \to x^*\Pi_x(A)$. Now $\Delta^*(f)$ corresponds to the composite

$$A \xrightarrow{\overline{f}} x^*\Pi_x(A) \cong \Pi_{\pi_1}\pi_1^*(A) \longrightarrow \Pi_{\pi_1}\Pi_\Delta\Delta^*\pi_2^*(A) \cong A$$

where the second factor is induced by the unit of $(\Delta^* \dashv \Pi_\Delta)$; but it is easily verified that the Beck–Chevalley isomorphism makes the diagram

$$\begin{tikzcd} x^*\Pi_x(A) \arrow[r, "\epsilon_A"] \arrow[d, "\cong"] & A \arrow[d, "\cong"] \\ \Pi_{\pi_1}\pi_2^*(A) \arrow[r] & \Pi_{\pi_1}\Pi_\Delta\Delta^*\pi_2^*(A) \end{tikzcd}$$

commute, where ϵ is the counit of $(x^* \dashv \Pi_x)$. Thus (A, f) satisfies the unit condition for an object of $\mathbf{Desc}(\mathbb{C}, (x))$ iff $(A, \overline{f})$ satisfies the unit condition for a $\mathbb{G}$-coalgebra. A similar diagram-chase shows that the composite $\pi_{23}^*(f) \circ \pi_{12}^*(f)$ and the morphism $\pi_{13}^*(f)$ in $\mathcal{C}^{J\times_I J\times_I J}$ transpose respectively to the composite

$$A \xrightarrow{\overline{f}} x^*\Pi_x(A) \xrightarrow{x^*\Pi_x(\overline{f})} x^*\Pi_x x^*\Pi_x(A)$$

and to

$$A \xrightarrow{\overline{f}} x^*\Pi_x(A) \xrightarrow{\delta_A} x^*\Pi_x x^*\Pi_x(A),$$

where δ is the comultiplication of $\mathbb{G}$; so (A, f) satisfies the cocycle condition iff $(A, \overline{f})$ satisfies the coassociativity condition for a coalgebra. Moreover, a morphism $h\colon A \to B$ between two objects equipped with such structures (A, f) and (B, g) clearly commutes in the appropriate sense with f and g iff it commutes with $\overline{f}$ and $\overline{g}$; so we have the required isomorphism of categories.

(ii) A further straightforward diagram-chase shows that this isomorphism identifies the functor $\theta_{(x)} : \mathcal{C}^I \to \mathbf{Desc}(\mathbb{C}, (x))$ with the comparison functor $\mathcal{C}^I \to (\mathcal{C}^J)_{\mathbb{G}}$; so this is immediate from the definitions.

(iii) If $\mathbb{C}$ is $\mathcal{S}$-complete, then x^* (has a right adjoint and) preserves equalizers; so (iii) follows immediately from the Crude (co)Monadicity Theorem (A1.1.2). □

Regarding descent morphisms for the canonical indexing, we have the following:

Proposition 1.5.6 *Let $\mathcal{S}$ be a cartesian category, and $x\colon J \to I$ a morphism of $\mathcal{S}$. Then*

(i) *x is a pre-descent morphism (for the canonical indexing of $\mathcal{S}$ over itself) iff it is a stable regular epimorphism (i.e. every pullback of x is a regular epimorphism).*

(ii) *In particular, if $\mathcal{S}$ is regular, then x is a pre-descent morphism iff it is a cover.*

(iii) *Suppose either that $\mathcal{S}$ is effective regular, or that $\mathcal{S}$ is cartesian and has pullback-stable coequalizers of reflexive pairs. Then every cover in $\mathcal{S}$ is a descent morphism.*

Proof (i) Let us write $(y, z)\colon K \rightrightarrows J$ for the kernel-pair of x, and w for the composite $xy = xz$. The assertion that x is a stable regular epimorphism is equivalent to saying that, for every $f\colon A \to I$, the top row of the diagram

$$\begin{tikzcd}
w^*(A) \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & x^*(A) \arrow[r] \arrow[d] & A \arrow[d, "f"] \\
K \arrow[r, shift left, "y"] \arrow[r, shift right, "z"'] & J \arrow[r, "x"] & I
\end{tikzcd}$$

(in which all the squares are pullbacks) is a coequalizer. Suppose this holds, and let $g\colon B \to I$ be another object of $\mathcal{S}/I$. Then the assertion that a morphism $x^*(A) \to x^*(B)$ over J is a morphism of $\mathbf{Desc}(\mathbb{S}, (x))$ (when its domain and

codomain are equipped with their canonical descent data) is equivalent to saying that there is a morphism $w^*(A) \to w^*(B)$ making both the squares

$$\begin{tikzcd} w^*(A) \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & x^*(A) \arrow[d] \\ w^*(B) \arrow[r, shift left] \arrow[r, shift right] & x^*(B) \end{tikzcd}$$

commute, and hence to saying that the composite $x^*(A) \to x^*(B) \to B$ factors uniquely through $x^*(A) \to A$. Thus x^* is full and faithful as a functor $\mathcal{S}/I \to$ $\mathbf{Desc}(\mathbb{S}, (x))$.

Conversely, suppose x is a pre-descent morphism. Then, for any $f\colon A \to I$ and any object B of $\mathcal{S}$, we see that morphisms $x^*(A) \to J^*(B) \cong x^*I^*(B)$ in $\mathbf{Desc}(\mathbb{S}, (x))$ correspond bijectively to morphisms $A \to I^*(B)$ in $\mathcal{S}/I$. Disentangling the definitions, this means that morphisms $x^*(A) \to B$ in $\mathcal{S}$ which have equal composites with $w^*(A) \rightrightarrows x^*(A)$ correspond bijectively to morphisms $A \to B$. So the top row of the diagram on the previous page is a coequalizer, i.e. x is a stable regular epimorphism.

(ii) follows immediately from (i) and A1.3.4.

(iii) First suppose $\mathcal{S}$ is effective regular. Let $(h\colon A \to J, f)$ be an object of $\mathbf{Desc}(\mathbb{S}, (x))$; then the projection $y^*(A) \to A$ and the composite $y^*(A) \to z^*(A) \to A$ define a pair of morphisms $y^*(A) \rightrightarrows A$, which is readily seen to be an equivalence relation on A. (Reflexivity and transitivity follow from the unit and cocycle conditions on the isomorphism $y^*(A) \to z^*(A)$; symmetry from the fact that it is an isomorphism and (y, z) is symmetric.) So we may form its coequalizer $q\colon A \to B$, and we may factor xh through q to obtain a morphism $k\colon B \to I$. We claim that the canonical morphism $A \to x^*(B)$ induced by q and h is an isomorphism. It is a cover, because the composite $y^*(A) \to A \to x^*(B)$ may be regarded as the pullback of $q\colon A \to B$ along x (where A is regarded as sitting over I by the morphism xh), and is thus a cover. And it is monic, because the pair (q, h) is jointly monic: if we have a pair $(l, m)\colon C \rightrightarrows A$ with $ql = qm$ and $hl = hm$, then we obtain a factorization $n\colon C \to y^*(A)$ of (l, m) through the kernel-pair $y^*(A) \rightrightarrows A$ of q, and then the composite $C \to y^*(A) \to K$ factors through the equalizer of $(y, z)\colon K \rightrightarrows J$ – but the latter is the diagonal map $J \to K$, so n factors through the corresponding pullback of $y^*(A)$, which forces $l = m$. Thus we have shown that every object of $\mathbf{Desc}(\mathbb{S}, (x))$ is isomorphic to one in the image of x^*: combining this with (ii), we have the desired result.

In the second case, it follows from A1.3.5 that $\mathcal{S}$ is regular; so the result follows immediately from (the dual of) 1.5.5(ii) and the Crude Monadicity Theorem (A1.1.2), since we saw in 1.4.7(i) that $\mathbb{S}$ has $\mathcal{S}$-indexed coproducts, and in A1.3.2(iii) that x is a cover iff $x^*\colon \mathcal{S}/I \to \mathcal{S}/J$ is conservative. □

Before proceeding further, we pause to note a result for non-regular categories which may be extracted from the proof of 1.5.6(iii), and which will be of use in Section C5.1:

Scholium 1.5.7 *Let $\mathcal{S}$ be a cartesian category, and let $\mathcal{P}$ be a class of morphisms of $\mathcal{S}$ such that*

(i) *$\mathcal{P}$ contains all isomorphisms, and is closed under composition;*

(ii) *$\mathcal{P}$ is stable under pullback;*

(iii) *every member of $\mathcal{P}$ is a coequalizer;*

(iv) *if $a, b\colon R \rightrightarrows A$ is an equivalence relation for which a (equivalently, b) is in $\mathcal{P}$, then it has a pullback-stable coequalizer which belongs to $\mathcal{P}$.*

Then every morphism of $\mathcal{P}$ is a descent morphism.

Proof Conditions (ii) and (iii) imply that the members of $\mathcal{P}$ are pre-descent morphisms, by 1.5.6(i). Suppose given a morphism $x\colon J \to I$ in $\mathcal{P}$, with kernel-pair $y, z\colon K \rightrightarrows J$, and an object $h\colon A \to J$ of $\mathcal{S}/J$ equipped with descent data relative to x. To simplify the notation, we shall use the descent data to identify the pullbacks of A along y and z, and denote their common value by A'. Thus we have a parallel pair $h^*(y), h^*(z)\colon A' \rightrightarrows A$, and as in the proof of 1.5.6 we know that this is an equivalence relation; moreover, $h^*(y)$ and $h^*(z)$ are in $\mathcal{P}$, since they are pullbacks of y and z, which are in turn pullbacks of x. So we may form their coequalizer $q\colon A \twoheadrightarrow B$, with the induced map $k\colon B \to I$. However, we do not have the information that $A' \rightrightarrows A$ is effective, so we cannot argue as in the proof of 1.5.6(iii) to show that $A \to x^*(B)$ is an isomorphism. Instead, we proceed as follows.

Let L denote the triple pullback $J \times_I J \times_I J$; then the cocycle condition for the descent data on A says that we may identify its pullbacks along the three projections $L \to I$, and we denote their common value by A''. Let u, v, w denote the three projections $L \to K$. Now form the pullback of the coequalizer $A' \rightrightarrows A \twoheadrightarrow B$ along $x^*(B) \to B$; since the latter is itself the pullback of x along k, and since $qk = xh$, it is easy to see that this has the form

$$\begin{tikzcd}
A'' \arrow[r, shift left, "h^*(v)"] \arrow[r, shift right, "h^*(w)"'] \arrow[d, "h^*(u)"] & A' \arrow[r, two heads] \arrow[d, "h^*(y)"] & x^*(B) \arrow[d] \\
A' \arrow[r, shift left, "h^*(y)"] \arrow[r, shift right, "h^*(z)"'] & A \arrow[r, two heads, "q"] & B
\end{tikzcd}$$

where the right-hand arrow on the top line is the composite of $h^*(z)$ with the comparison map $A \to x^*(B)$. But it is also easy to see that

$$\begin{tikzcd} A'' \arrow[r, shift left, "h^*(v)"] \arrow[r, shift right, "h^*(w)"'] & A' \arrow[r, "h^*(z)"] & A \end{tikzcd}$$

has the structure of a split coequalizer, with splittings induced by pulling back the diagonal map $d\colon J \to K$ and $1 \times d\colon K \to L$ along h. Hence $A \to x^*(B)$ must be an isomorphism. $\square$

We now turn to the coherent coverage on a positive coherent category $\mathcal{S}$. If $(x_i\colon J_i \to I \mid 1 \le i \le n)$ is a covering family for this coverage, we may 'factor' it as the composite of the family of coprojections $(\nu_i\colon J_i \to J \mid 1 \le i \le n)$ (where J is the coproduct of the J_i) and the singleton cover $(x\colon J \to I)$ induced by the x_i. Clearly, an $\mathcal{S}$-indexed category is a (pre-)stack for the coherent coverage iff it satisfies the (pre-)descent condition for families of these two kinds. But the singleton families are covering for the regular coverage, so we have already dealt with these in 1.5.5 and 1.5.6; the other families consist of monomorphisms, and here again we may simplify the definition of descent data. Since a morphism $x\colon J \to I$ is monic iff

$$\begin{tikzcd} J \arrow[r, "1"] \arrow[d, "1"'] & J \arrow[d, "x"] \\ J \arrow[r, "x"] & I \end{tikzcd}$$

is a pullback, the 'unit' condition (*a*) in the definition of descent data on a family $(A_i \in \mathrm{ob}\, \mathcal{C}^{J_i} \mid 1 \le i \le n)$ (relative to a family of monomorphisms $(x_i\colon J_i \rightarrowtail J)$) becomes simply the assertion that the 'diagonal' components f_{ii} of the data are (modulo canonical isomorphisms) simply the identity morphisms on the A_i. So we can forget about these components, and simply describe the descent data as a family of *isomorphisms* $(f_{ij}\colon y_{ij}^*(A_i) \to z_{ij}^* A_j \mid i < j)$, subject to the condition that

$$\begin{tikzcd} t_{ijk}^* y_{ij}^*(A_i) \arrow[r, "t_{ijk}^*(f_{ij})"] \arrow[d] & t_{ijk}^* z_{ij}^*(A_j) \arrow[r] & v_{ijk}^* y_{jk}^*(A_j) \arrow[d, "v_{ijk}^*(f_{jk})"] \\ u_{ijk}^* y_{ik}^*(A_i) \arrow[r, "u_{ijk}^*(f_{ik})"] & u_{ijk}^* z_{ik}^*(A_k) \arrow[r] & v_{ijk}^* z_{jk}^*(A_k) \end{tikzcd}$$

commutes for all $i < j < k$. (Here we need to assume explicitly that the f_{ij} are isomorphisms, because we have suppressed mention of the f_{ji} which are their inverses.)

In practice, we may usually restrict our attention still further, to families consisting of just two monomorphisms (plus the empty family), since more general finite families can be built up from these by induction. For a family $(x_1 : J_1 \rightarrowtail J, x_2 : J_2 \rightarrowtail J)$ the 'cocycle' condition above becomes vacuous, and we have

Lemma 1.5.8 *An $\mathcal{S}$-indexed category $\mathbb{C}$ satisfies the descent condition for a pair of monomorphisms $(x_1 : J_1 \rightarrowtail J, x_2 : J_2 \rightarrowtail J)$ in $\mathcal{S}$ iff the diagram*

$$\begin{array}{ccc} C^J & \xrightarrow{x_1^*} & C^{J_1} \\ \downarrow{\scriptstyle x_2^*} & & \downarrow{\scriptstyle y_{12}^*} \\ C^{J_2} & \xrightarrow{z_{12}^*} & C^{J_1 \cap J_2} \end{array}$$

is a (pseudo-)pullback in $\mathfrak{CAT}$. □

Corollary 1.5.9 *Let $\mathcal{S}$ be a positive coherent category. Then a nonempty $\mathcal{S}$-indexed category $\mathbb{C}$ is a (pre-)stack for the coherent coverage on $\mathcal{S}$ iff it is a (pre-)stack for the regular coverage, and in addition we have $C^0 \simeq \mathbf{1}$ and, for each pair of objects (I, J) the functor $(\nu_1^*, \nu_2^*) : C^{I \amalg J} \to C^I \times C^J$ is an equivalence (resp. full and faithful).*

Proof We have already observed that covering families can be decomposed into finite covering families of (pairwise disjoint) monomorphisms and regular covers. The initial object 0 of $\mathcal{S}$ is covered by the empty family, and the descent condition for this family says that $C^0 \simeq \mathbf{1}$. (The pre-descent condition for it says that C^0 is either empty or equivalent to $\mathbf{1}$; but if any fibre C^I is nonempty then C^0 must be nonempty.) Hence, for a disjoint pair of monomorphisms, the pullback diagram of 1.5.8 reduces to a product diagram. □

As before, we investigate what these conditions mean for the canonical indexing of $\mathcal{S}$ over itself.

Lemma 1.5.10

(i) *If $\mathcal{S}$ is a coherent category, then $\mathbb{S}$ is a pre-stack for the coverage whose covering families are finite families of monomorphisms*

$$(x_i : J_i \rightarrowtail J \mid 1 \leq i \leq n)$$

in $\mathcal{S}$ such that $J = \bigcup_{i=1}^n J_i$.

(ii) *If $\mathcal{S}$ is a pretopos, then $\mathbb{S}$ is a stack for the coherent coverage on $\mathcal{S}$.*

Proof (i) It suffices to consider the empty family and families with two members. For the first, we use the fact that the initial object of $\mathcal{S}$ is strict (A1.4.1). For doubleton families, we use an argument much like that of 1.5.6(i), relying on A1.4.3 and the stability of binary unions under pullback in $\mathcal{S}$.

(ii) The reason why we get a pre-stack rather than a stack in (i) is that we cannot, in a general coherent category, form the pushout of a pair of monomorphisms with common domain. But in a pretopos we can do so (A1.4.8), and so in this case $\mathbb{S}$ is a stack for the coverage described in (i), by an argument like that of 1.5.6(iii). Since it is also a stack for the regular coverage by 1.5.6(iii), the result follows. $\square$

We remark in passing that a *finitary disjunctive category* (also called a lextensive category; cf. [227]) is one with finite limits, a strict initial object and binary coproducts which are disjoint and stable under pullback. Equivalently, it is a cartesian category $\mathcal{S}$ with finite coproducts, such that $\mathbb{S}$ is a stack for the coverage consisting of finite families of coprojections. We shall meet disjunctive categories again, albeit briefly, in Part D (see D1.3.6).

Suggestions for further reading: Bourn [171], Bunge & Paré [205], Giraud [406, 407], Grothendieck [421], Reiterman & Tholen [1003].

B2

INTERNAL AND LOCALLY INTERNAL CATEGORIES

B2.1 Review of enriched categories

The theory of enriched categories is normally studied in the context of a base category $\mathcal{V}$ with a monoidal structure (usually, but not always, symmetric). However, we are primarily interested in studying categories enriched in a topos, where the monoidal structure is taken to be the cartesian one (i.e. that induced by finite products), and so we shall save ourselves a bit of extra work by developing the theory only for cartesian monoidal structures. Throughout this section, therefore, our base category $\mathcal{V}$ will be assumed to have finite products.

Definition 2.1.1 (*a*) A *$\mathcal{V}$-category* (or *$\mathcal{V}$-enriched category*) $\mathcal{C}$ consists of a collection of *objects* together with

(i) for each pair of objects (A, B), an object $\mathcal{C}(A, B)$ of $\mathcal{V}$;

(ii) for each object A, a morphism $1_A : 1 \to \mathcal{C}(A, A)$ in $\mathcal{V}$;

(iii) for each triple of objects (A, B, C), a morphism

$$c_{A,B,C} : \mathcal{C}(B, C) \times \mathcal{C}(A, B) \longrightarrow \mathcal{C}(A, C)$$

in $\mathcal{V}$,

such that the $\mathrm{c}_{A,B,C}$ are associative in the sense that

$$\begin{tikzcd}
\mathcal{C}(C,D) \times \mathcal{C}(B,C) \times \mathcal{C}(A,B) \arrow[r, "1 \times \mathrm{c}"] \arrow[d, "\mathrm{c} \times 1"] & \mathcal{C}(C,D) \times \mathcal{C}(A,C) \arrow[d, "c"] \\
\mathcal{C}(B,D) \times \mathcal{C}(A,B) \arrow[r, "c"] & \mathcal{C}(A,D)
\end{tikzcd}$$

commutes for each quadruple (A, B, C, D), and the 1_A act as identities for c in a similar sense.

(*b*) If $\mathcal{C}$ and $\mathcal{D}$ are $\mathcal{V}$-categories, a *$\mathcal{V}$-functor* $F : \mathcal{C} \to \mathcal{D}$ consists of a function $A \mapsto FA$ from objects of $\mathcal{C}$ to objects of $\mathcal{D}$ and, for each pair of objects (A, B), a morphism $F_{A,B} : \mathcal{C}(A, B) \to \mathcal{D}(FA, FB)$, such that the $F_{A,B}$ are compatible

in the obvious sense with composition and identities. We call F a $\mathcal{V}$*-embedding* (resp. a $\mathcal{V}$*-full embedding*) if each $F_{A,B}$ is monic (resp. an isomorphism) in $\mathcal{V}$.

(c) If F and $G\colon \mathcal{C} \to \mathcal{D}$ are $\mathcal{V}$-functors, a $\mathcal{V}$*-natural transformation* $\alpha\colon F \to G$ is a function assigning to each object A of $\mathcal{C}$ a morphism $\alpha_A\colon 1 \to \mathcal{D}(FA, GA)$, such that the diagram

$$\begin{tikzcd}
\mathcal{C}(A,B) \arrow[r, "\alpha_B \times F_{A,B}"] \arrow[d, "G_{A,B} \times \alpha_A"'] & \mathcal{D}(FB,GB) \times \mathcal{D}(FA,FB) \arrow[d, "c_{FA,FB,GB}"] \\
\mathcal{D}(GA,GB) \times \mathcal{D}(FA,GA) \arrow[r, "c_{FA,GA,GB}"] & \mathcal{D}(FA,GB)
\end{tikzcd}$$

commutes for each pair of objects (A, B).

We write $\mathcal{V}$-$\mathfrak{Cat}$ for the 2-category of $\mathcal{V}$-enriched categories, functors and natural transformations.

If $\mathcal{C}$ is a $\mathcal{V}$-category, the *underlying ordinary category* $|\mathcal{C}|$ of $\mathcal{C}$ has the same objects as $\mathcal{C}$, and its morphisms $A \to B$ are defined to be morphisms $1 \to \mathcal{C}(A, B)$ in $\mathcal{V}$, with composition defined by $gf = c_{A,B,C}(g, f)$. We similarly define underlying ordinary functors and natural transformations of $\mathcal{V}$-enriched ones; note in particular that the underlying ordinary functor of a $\mathcal{V}$-embedding (resp. a $\mathcal{V}$-full embedding) is faithful (resp. full and faithful). We refer to $\mathcal{C}$ as a $\mathcal{V}$*-enrichment* of the ordinary category $|\mathcal{C}|$.

In the case when $\mathcal{V} = \mathbf{Set}$, a $\mathcal{V}$-enriched category is just a locally small category; and when $\mathcal{V} = \mathbf{Cat}$, a $\mathcal{V}$-enriched category is a locally small 2-category. In the former case, a given ordinary category has at most one $\mathcal{V}$-enrichment (up to canonical isomorphism); this is because the terminal object 1 is a generator for **Set**. In general, however, there may be many different ways of enriching a given category over a given $\mathcal{V}$. For example, the ordinary category **Cat** has (at least) three different enrichments in itself: for the first we take $\mathbf{Cat}(\mathcal{C}, \mathcal{D})$ to be the usual functor category $[\mathcal{C}, \mathcal{D}]$, for the second we take the subcategory of all functors $\mathcal{C} \to \mathcal{D}$ and natural isomorphisms between them, and for the third we take the discrete category of functors $\mathcal{C} \to \mathcal{D}$. Similar remarks apply to enrichments of ordinary functors: any functor between locally small categories admits a unique **Set**-enrichment, but (for example) the ordinary functor $(-)^{\mathrm{op}}\colon \mathbf{Cat} \to \mathbf{Cat}$ is **Cat**-enrichable with respect to the second and third **Cat**-enrichments of **Cat** defined above, but not with respect to the first.

Lemma 2.1.2 *A cartesian closed category $\mathcal{V}$ has a canonical enrichment over itself.*

Proof We define $\mathcal{V}(A, B)$ to be the exponential B^A, 1_A to be the transpose of $\pi_2 \colon 1 \times A \to A$, and $c_{A,B,C}$ to be the transpose of

$$C^B \times B^A \times A \xrightarrow{1 \times \mathrm{ev}} C^B \times B \xrightarrow{\mathrm{ev}} C\,.$$

The remaining details are straightforward. □

For the reasons indicated above, we cannot hope to have any sort of converse to 2.1.2; an enrichment of $\mathcal{V}$ over itself need not derive from a cartesian closed structure. In contrast, we shall see in the next section that an *indexed* enrichment of a category over itself necessarily derives from a locally cartesian closed structure.

If $\mathcal{V}$ is a cartesian closed category and $\mathcal{C}$ is a $\mathcal{V}$-category, then for each object A of $\mathcal{C}$ we have a $\mathcal{V}$-functor $\mathcal{C}(A, -) \colon \mathcal{C} \to \mathcal{V}$ sending B to $\mathcal{C}(A, B)$; $\mathcal{C}(A, -)_{B,C} \colon \mathcal{C}(B, C) \to \mathcal{C}(A, C)^{\mathcal{C}(A,B)}$ is of course just the exponential transpose of the composition map $c_{A,B,C}$ of $\mathcal{C}$. We have an enriched version of the Yoneda lemma:

Lemma 2.1.3 *Let $\mathcal{V}$ be a cartesian closed category, $\mathcal{C}$ a $\mathcal{V}$-category, $F \colon \mathcal{C} \to \mathcal{V}$ a $\mathcal{V}$-functor and A an object of $\mathcal{C}$. Then there is a bijection between $\mathcal{V}$-natural transformations $\mathcal{C}(A, -) \to F$ and morphisms $1 \to FA$ in $\mathcal{C}$.*

Proof Given a $\mathcal{V}$-natural transformation α, we have the composite

$$1 \xrightarrow{1_A} \mathcal{C}(A, A) \xrightarrow{\alpha_A} FA;$$

conversely, given a morphism $x \colon 1 \to FA$, we have the natural transformation whose B-component is

$$\mathcal{C}(A, B) \xrightarrow{F_{A,B} \times x} FB^{FA} \times FA \xrightarrow{\mathrm{ev}} FB\,.$$

The verification that these constructions are inverse to each other is straightforward. □

If $\mathcal{C}$ is a small $\mathcal{V}$-category (i.e. one with a set of objects) and $\mathcal{V}$ is complete, we may make the $\mathcal{V}$-functors $\mathcal{C} \to \mathcal{V}$ into the objects of a $\mathcal{V}$-category $[\mathcal{C}, \mathcal{V}]$; and then the Yoneda lemma yields a $\mathcal{V}$-full embedding $\mathcal{C}^{\mathrm{op}} \to [\mathcal{C}, \mathcal{V}]$. However, we shall not need to develop this aspect of enriched category theory in the present work.

The following characterization of enriched functors from a cartesian closed category to itself is often useful.

Lemma 2.1.4 *Let $\mathcal{V}$ be a cartesian closed category, and let $F \colon \mathcal{V} \to \mathcal{V}$ be a functor. Then specifying a $\mathcal{V}$-enrichment of F is equivalent to specifying a natural*

transformation

$$U \times FV \xrightarrow{\tau_{U,V}} F(U \times V)$$

(between functors $\mathcal{V} \times \mathcal{V} \to \mathcal{V}$), such that $\tau_{1,V}$ is the canonical isomorphism $1 \times FV \cong FV \cong F(1 \times V)$ and the diagram

$$\begin{tikzcd}
U \times V \times FW \arrow[r, "1_U \times \tau_{V,W}"] \arrow[dr, "\tau_{U \times V,W}"'] & U \times F(V \times W) \arrow[d, "\tau_{U,V \times W}"] \\
& F(U \times V \times W)
\end{tikzcd}$$

commutes for all U, V, W.

Proof Given an enrichment $F_{U,V} \colon V^U \to FV^{FU}$ of F, we define $\tau_{U,V}$ to be exponential transpose of the composite

$$U \xrightarrow{\lambda} (U \times V)^V \xrightarrow{F_{V,U \times V}} F(U \times V)^{FV}$$

where λ is the unit of the exponential adjunction; conversely, given τ, we define $F_{U,V}$ to be the transpose of

$$V^U \times FU \xrightarrow{\tau_{V^U,U}} F(V^U \times U) \xrightarrow{F(\mathrm{ev})} FV .$$

A straightforward diagram-chase shows that the two conditions on τ are equivalent to the compatibility of the $F_{U,V}$ with identity morphisms and composition, and that the two constructions above are inverse to each other. The fact that $F_{U,V}$ is an enrichment of the original functor F (i.e. that the effect of composing it with morphisms $1 \to V^U$ corresponds to the effect of F on morphisms $U \to V$) is built into the naturality of τ in its first variable. □

Enriched functors from a cartesian closed category to itself are sometimes called *strong functors*. If $F \colon \mathcal{V} \to \mathcal{V}$ is a functor from a category with finite products to itself, a natural transformation τ satisfying the conditions of 2.1.4 is called a *strength* for F, even if $\mathcal{V}$ is not cartesian closed.

Example 2.1.5 Suppose $\mathcal{V}$ has list objects in the sense of A2.5.15. Then the assignment $U \mapsto LU$ can be made into a functor $\mathcal{V} \to \mathcal{V}$, as we observed there; and this functor has a canonical strength $\tau_{U,V} \colon U \times LV \to L(U \times V)$, namely

the unique morphism making

$$\begin{tikzcd}
U \arrow[r, "1_U \times o_V"] \arrow[d] & U \times LV \arrow[d, "\tau_{U,V}"] & U \times V \times LV \arrow[l, "1_U \times s_V"'] \arrow[d, "\pi_1, \pi_2, \tau_{U,V}\pi_3"] \\
1 \arrow[r, "o_{U\times V}"] & L(U \times V) & U \times V \times L(U \times V) \arrow[l, "s_{U\times V}"']
\end{tikzcd}$$

commute. (Intuitively, $\tau_{U,V}$ is the operation which takes an element of U and a list of elements of V, and produces a list of elements of $U \times V$ whose first coordinates are all equal to the given element of U.) It is straightforward to verify that $\tau_{U,V}$ satisfies the conditions of 2.1.4. In the case when $\mathcal{V}$ is a topos with a natural number object, the fact that L is a $\mathcal{V}$-enriched functor may alternatively be deduced from the fact that it is a $\mathcal{V}$-indexed functor (i.e. the slice categories $\mathcal{V}/U$ all have list objects, and pullback functors preserve them), as we shall see in the next section (cf. 2.2.2).

It is also not hard to see that, for an ordinary natural transformation $\alpha\colon F \to G$ between strong functors $\mathcal{V} \to \mathcal{V}$, the condition that α should be $\mathcal{V}$-enriched (i.e. that it should make the diagram of 2.1.1(c) commute) is equivalent to requiring that it should commute with the strengths of F and G, in the sense that

$$\begin{tikzcd}
U \times FV \arrow[r, "\tau_{U,V}"] \arrow[d, "1_U \times \alpha_V"] & F(U \times V) \arrow[d, "\alpha_{U\times V}"] \\
U \times GV \arrow[r, "\tau'_{U,V}"] & G(U \times V)
\end{tikzcd}$$

commutes. Of course, a natural transformation with this property is called a *strong natural transformation.* The following result can thus be stated for categories with finite products, since its proof makes no use of the closed structure – although, in the applications, we shall always use it for cartesian closed categories.

Lemma 2.1.6 *Let $\mathcal{V}$ be a category with finite products, and $T\colon \mathcal{V} \to \mathcal{V}$ an (ordinary) functor which preserves finite products. Suppose further that there is a natural transformation $\eta\colon 1_{\mathcal{V}} \to T$. Then there is a unique strength on T making η into a strong natural transformation. If η is the unit of a monad structure $\mathbb{T} = (T, \eta, \mu)$, then this strength also makes μ a strong natural transformation.*

Proof We define $\tau_{U,V}$ to be the composite

$$U \times TV \xrightarrow{\eta_U \times 1_V} TU \times TV \cong T(U \times V);$$

it is clear that this is natural in U and V, and that it satisfies the first condition of 2.1.4. For the second, we need to observe that

$$\begin{tikzcd} U \times V \arrow[r, "\eta_{U \times V}"] \arrow[dr, "\eta_U \times \eta_V"'] & T(U \times V) \arrow[d, "\cong"] \\ & TU \times TV \end{tikzcd}$$

commutes; but this is immediate from the naturality of η, since the vertical isomorphism is obtained by applying T to the two projections from $U \times V$. The same commutative diagram verifies that η is a strong natural transformation $1_{\mathcal{V}} \to T$. To verify that μ, if it exists, is also strong, we need the similar diagram identifying $\mu_{U \times V}$ with $\mu_U \times \mu_V$, plus the monad identity $\mu_U \cdot T\eta_U = 1_{TU}$.

For the uniqueness, suppose $\sigma_{U,V} \colon U \times TV \to T(U \times V)$ is any strength for T. Composing it with $T\pi_1$ and $T\pi_2$, we get natural transformations $\alpha_{U,V} \colon U \times TV \to TU$ and $\beta_{U,V} \colon U \times TV \to TV$, which together determine σ. But $\beta_{U,V} = \beta_{1,V} \circ \pi_2$ by naturality, and $\beta_{1,V} = 1_{TV}$ by the first condition of 2.1.4. Similarly, $\alpha_{U,V} = \alpha_{U,1} \circ \pi_1$, and $\alpha_{U,1} = \eta_U$ by the fact that η is strong. So we have identified σ with the strength τ defined above. □

Remark 2.1.7 If $\mathcal{V}$ is a category with finite products, and $T \colon \mathcal{V} \to \mathcal{V}$ is a strong functor which preserves 1, then there is a unique strong natural transformation from the identity functor to T, namely that consisting of the components $\tau_{(-),1}$ of the strength of T. (For cartesian closed $\mathcal{V}$, this is a particular case of 2.1.3, since we can identify the identity functor with $\mathcal{V}(1, -)$.) Thus 2.1.6 tells us that, for an ordinary functor T which preserves finite products, there is a bijection between the possible strengths for T and the (ordinary) natural transformations $1_{\mathcal{V}} \to T$.

If $F \colon \mathcal{V} \to \mathcal{W}$ is a functor preserving finite products, then for any $\mathcal{V}$-enriched category $\mathcal{C}$ we may define a $\mathcal{W}$-enriched category $F_\bullet \mathcal{C}$ having the same objects as $\mathcal{C}$, but

$$F_\bullet \mathcal{C}(A, B) = F(\mathcal{C}(A, B)) \ ,$$

and composition and identity operations obtained by applying F to those of $\mathcal{C}$. In fact it is not hard to see that $F_\bullet$ is a 2-functor $\mathcal{V}\text{-}\mathfrak{Cat} \to \mathcal{W}\text{-}\mathfrak{Cat}$. Note, however, that $\mathcal{C}$ and $F_\bullet \mathcal{C}$ do not in general have the same underlying ordinary category, unless *either* F happens to be full and faithful, *or* F has a left adjoint which preserves 1. (The latter case is of importance for us, since it applies when F is the direct image of a geometric morphism.)

If $F: \mathcal{V} \to \mathcal{W}$ is a finite-product-preserving functor between cartesian closed categories, then the comparison maps $\theta_{U,V}: F(V^U) \to FV^{FU}$, defined before A1.5.2, give rise to a $\mathcal{W}$-functor (which we shall also denote by F) $F_\bullet \mathcal{V} \to \mathcal{W}$. We note that this functor is a $\mathcal{W}$-full embedding iff F is a cartesian closed functor (and it is a $\mathcal{W}$-embedding iff F is sub-cartesian-closed, as defined before C3.1.1). In particular, we have

Lemma 2.1.8 *If $\mathcal{V}$ is a locally cartesian closed category, then for any $x: U \to V$ in $\mathcal{V}$ the pullback functor x^* induces a $\mathcal{V}/U$-full embedding $(x^*)_\bullet(\mathcal{V}/V) \to \mathcal{V}/U$.*

Proof We saw in A1.5.2 that pullback functors are always cartesian closed. □

The above result is an important part of the motivation for the definition of a locally internal category at the start of the next section.

Suggestions for further reading: Kelly [578], Kock [603, 604].

B2.2 Locally internal categories

In this section we introduce the notion of locally internal category, due to J. Penon [955], which synthesizes the concepts of indexed category and enriched category. Since the definition involves enrichments in all the slice categories $\mathcal{S}/I$ of our base category $\mathcal{S}$, we must henceforth require the latter to have all finite limits.

Definition 2.2.1 Let $\mathcal{S}$ be a cartesian category. A *locally internal category* $\mathbb{C}$ over $\mathcal{S}$ assigns to each object I of $\mathcal{S}$ an $\mathcal{S}/I$-enriched category $\mathcal{C}^I$, to each morphism $x: I \to J$ in $\mathcal{S}$ an $\mathcal{S}/I$-enriched full embedding $x^*: (x^*)_\bullet(\mathcal{C}^J) \to \mathcal{C}^I$, and to each composable pair $(x: I \to J, y: J \to K)$ an $\mathcal{S}/I$-enriched natural isomorphism $\theta_{x,y}$ as in

$$\begin{tikzcd}
(x^*)_\bullet(y^*)_\bullet\mathcal{C}^K \arrow[r, "(x^*)_\bullet(y^*)"] \arrow[d] \arrow[dr, phantom, "\downarrow \theta_{x,y}"] & (x^*)_\bullet\mathcal{C}^J \arrow[d, "x^*"] \\
((yx)^*)_\bullet\mathcal{C}^K \arrow[r, "(yx)^*"'] & \mathcal{C}^I
\end{tikzcd}$$

(where the left vertical arrow is induced by the canonical isomorphism $x^*y^* \cong (yx)^*$), subject to the usual coherence conditions. *Locally internal functors* and *locally internal natural transformations* are defined in the obvious way; they form the 1-cells and 2-cells of the 2-category of locally internal categories over $\mathcal{S}$, which we denote by $\mathfrak{Cat}_\mathcal{S}$.

It is clear that, just as an enriched category has an underlying ordinary category, so a locally internal category over $\mathcal{S}$ has an underlying $\mathcal{S}$-indexed category. What is, perhaps, more surprising is that a locally internal category is determined up to isomorphism by its underlying indexed category.

Theorem 2.2.2 *For any cartesian category $\mathcal{S}$, $\mathfrak{Cat}_{\mathcal{S}}$ is strongly equivalent to the full sub-2-category of $\mathfrak{CAT}_{\mathcal{S}}$ whose objects are locally small $\mathcal{S}$-indexed categories, as defined in* 1.3.12.

Proof First suppose given a locally internal category $\mathbb{C}$. If A and B are objects of $\mathcal{C}^J$ and $\mathcal{C}^K$ respectively, let $(x,y)\colon I \to J \times K$ be the object $\mathcal{C}^{J\times K}(\pi_1^*A, \pi_2^*B)$ of $\mathcal{S}/J \times K$. Then, for any object $(z,w)\colon L \to J \times K$ of $\mathcal{S}/J \times K$, morphisms $u\colon L \to I$ over $J\times K$ correspond to morphisms $1_L \to (z,w)^*(\mathcal{C}^{J\times K}(\pi_1^*A, \pi_2^*B)) \cong \mathcal{C}^L(z^*A, w^*B)$ in $\mathcal{S}/L$; that is, to morphisms $z^*A \to w^*B$ in the underlying ordinary category of $\mathcal{C}^L$. Thus the underlying indexed category of $\mathbb{C}$ is locally small.

Conversely, let $\mathbb{C}$ be a locally small indexed category. Given any two objects A, B of the same fibre $\mathcal{C}^I$, we have an object $J \to I \times I$ indexing morphisms from A to B, as in 1.3.12; we define $\mathcal{C}^I(A,B)$ to be the pullback of this object along the diagonal $\Delta\colon I \to I \times I$. Then $\mathcal{C}^I(A,B)$ represents the functor on $(\mathcal{S}/I)^{\mathrm{op}}$ which sends $x\colon K \to I$ to the class of morphisms $x^*A \to x^*B$ in $\mathcal{C}^K$; using the Yoneda lemma it is straightforward to define morphisms $1_I \to \mathcal{C}^I(A,A)$ and $\mathcal{C}^I(B,C) \times \mathcal{C}^I(A,B) \to \mathcal{C}^I(A,C)$ giving $\mathcal{C}^I$ the structure of an $(\mathcal{S}/I)$-enriched category. Further, we see that for any $x\colon K \to I$ the objects $\mathcal{C}^K(x^*A, x^*B)$ and $x^*(\mathcal{C}^I(A,B))$ represent the same functor on $(\mathcal{S}/K)^{\mathrm{op}}$, so they are canonically isomorphic; thus $x^*\colon \mathcal{C}^I \to \mathcal{C}^K$ becomes an $(\mathcal{S}/K)$-enriched full embedding. So $\mathbb{C}$ can be enriched to a locally internal category.

Further, this is the only possible locally internal structure on $\mathbb{C}$ (up to canonical isomorphism); for, if we have such a structure, then for any $x\colon K \to I$ morphisms $x \to \mathcal{C}^I(A,B)$ in $\mathcal{S}/I$ correspond to morphisms $1_K \to \mathcal{C}^K(x^*A, x^*B)$ in $\mathcal{S}/K$, i.e. to morphisms $x^*A \to x^*B$ in $\mathcal{C}^K$. So $\mathcal{C}^I(A,B)$, as the representation of a functor on $(\mathcal{S}/I)^{\mathrm{op}}$, is determined up to unique isomorphism for each pair of objects (A,B) of $\mathcal{C}^I$.

A similar argument shows that any indexed functor $F\colon \mathbb{C} \to \mathbb{D}$ between locally small indexed categories admits a unique enrichment to a locally internal functor: given objects A and B of $\mathcal{C}^I$, we have for any $x\colon K \to I$ a mapping induced by F^K from morphisms $x^*A \to x^*B$ in $\mathcal{C}^K$ to morphisms

$$x^*F^IA \cong F^Kx^*A \longrightarrow F^Kx^*B \cong x^*F^IB$$

in $\mathcal{D}^K$, and so there is a canonical morphism $(F^I)_{A,B} : \mathcal{C}^I(A,B) \to \mathcal{D}^I(F^IA, F^IB)$. This is easily seen to be an enrichment of F^I, and to be the unique such enrichment which is compatible with the transition functors x^* of $\mathbb{C}$ and $\mathbb{D}$. In the same way, any indexed natural transformation between locally internal functors is automatically enriched. So the operation of equipping locally small

indexed categories, and indexed functors and natural transformations between them, with the indicated enrichments defines a functor from the indicated sub-2-category of $\mathfrak{CAT}_{\mathcal{S}}$ to $\mathfrak{Cat}_{\mathcal{S}}$, which is inverse to the forgetful functor up to natural isomorphism. □

From now on, we shall use the terms 'locally internal category' and 'locally small indexed category' interchangeably, and we shall not bother to distinguish notationally between the two concepts.

Next, we establish a result promised in Section B1.3.

Lemma 2.2.3 *Let $\mathcal{S}$ be a cartesian category. Then the canonical indexing of $\mathcal{S}$ over itself is locally small iff $\mathcal{S}$ is locally cartesian closed.*

Proof If $\mathcal{S}$ is locally cartesian closed, then by 2.1.8 $\mathbb{S}$ is the underlying indexed category of a locally internal category, and therefore locally small. Conversely, suppose $\mathbb{S}$ is locally small; let $x\colon J \to I$ and $y\colon K \to I$ be two objects of $\mathcal{S}/I$. Then, for any $z\colon L \to I$, morphisms $z \to (\mathcal{S}/I)\,(x,y)$ in $\mathcal{S}/I$ correspond to morphisms $z^*x \to z^*y$ in $\mathcal{S}/L$, and hence to morphisms $x \times z \cong \Sigma_z z^* x \to y$ in $\mathcal{S}/I$; that is, $(\mathcal{S}/I)\,(x,y)$ has the universal property of an exponential y^x in $\mathcal{S}/I$. □

Generalizing one direction of 2.2.3, we have

Lemma 2.2.4 *Suppose $\mathcal{S}$ is locally cartesian closed. Then, for any locally small $\mathcal{S}$-indexed category $\mathbb{C}$, the free category with $\mathcal{S}$-indexed coproducts* Fam($\mathbb{C}$) *(cf.* 1.4.16*) is locally small.*

Proof We recall that objects of $\mathrm{Fam}(\mathcal{C})^I$ are pairs (x, A), where $x\colon J \to I$ in $\mathcal{S}$ and A is an object of $\mathcal{C}^J$. Suppose given two such objects (x_1, A_1) and (x_2, A_2) in the fibres over I_1 and I_2 respectively: let $K \to I_1 \times I_2$ denote the exponential $(\pi_2^*(x_2))^{(\pi_1^*(x_1))}$ in $\mathcal{S}/I_1 \times I_2$, and form the diagram

$$\begin{tikzcd}
I_1 \times J_2 & L \arrow[l, "e"'] \arrow[r, "k"] \arrow[d, "l"] & K \arrow[d] \\
& J_1 \times I_2 \arrow[r, "x_1 \times 1"] & I_1 \times I_2
\end{tikzcd}$$

where the square is a pullback and e is the evaluation map associated with the exponential. Now let $M \to L$ be the object indexing morphisms $l^*\pi_1^*(A_1) \to e^*\pi_2^*(A_2)$ in the fibre $\mathcal{C}^L$; then we claim that the composite $\Pi_k(M) \to K \to I_1 \times I_2$ indexes morphisms $(x_1, A_1) \to (x_2, A_2)$ in Fam($\mathbb{C}$). For, given any object $(z_1, z_2)\colon Z \to I_1 \times I_2$ and a morphism $Z \to \Pi_k(M)$ over $I_1 \times I_2$, the composite $Z \to \Pi_k(M) \to K$ defines a morphism $y\colon z_1^*(x_1) \to z_2^*(x_2)$ in $\mathcal{S}/Z$, and then the factorization of this through $\Pi_k(M)$ defines a morphism $z_1^*(x_1, A_1) \to z_2^*(x_2, A_2)$ in $\mathrm{Fam}(\mathcal{C})^Z$. □

If $\mathcal{S}$ is a topos with a natural number object, then the same argument will show that the construction $\mathrm{Fam}_f(-)$ of 1.4.17 preserves local smallness.

Lemma 2.2.5 *Let $\mathcal{S}$ and $\mathcal{T}$ be cartesian categories, and $F\colon \mathcal{S} \to \mathcal{T}$ a functor. If the $\mathcal{S}$-indexed category $I \mapsto \mathcal{T}/FI$ of 1.2.2(d) is locally small, then F has a right adjoint.*

Proof First suppose F preserves the terminal object. Given an object A of $\mathcal{T} \cong \mathcal{T}/1$, we define RA to be the object $\mathcal{T}(1, A)$ of $\mathcal{S}$. Then, for any object I of $\mathcal{S}$, morphisms $I \to RA$ in $\mathcal{S}$ correspond to morphisms $FI^*1 \to FI^*A$ in $\mathcal{T}/FI$, and hence to morphisms $FI \to A$ in $\mathcal{T}$. This correspondence is clearly natural in I and A; so R is right adjoint to F.

In the general case, we can as usual factor F as

$$\mathcal{S} \xrightarrow{\hat{F}} \mathcal{T}/F1 \xrightarrow{\Sigma_{F1}} \mathcal{T}$$

where $\hat{F}$ preserves the terminal object; and the argument above shows that $\hat{F}$ has a right adjoint. But Σ_{F1} also has a right adjoint, namely $(F1)^*$; so F has a right adjoint. □

Corollary 2.2.6 *Let $F\colon \mathcal{S} \to \mathcal{T}$ be a functor between cartesian categories, and suppose $\mathcal{T}$ is locally cartesian closed. Then the following assertions are equivalent:*

(i) *F has a right adjoint.*

(ii) *$F^*\colon \mathfrak{CAT}_{\mathcal{T}} \to \mathfrak{CAT}_{\mathcal{S}}$ preserves local smallness.*

(iii) *The indexing of $\mathcal{T}$ over $\mathcal{S}$ induced by F is locally small.*

Proof (i) $\Rightarrow$ (ii) is a special case of 1.3.17; (ii) $\Rightarrow$ (iii) since the canonical indexing of $\mathcal{T}$ over itself is locally small, by 2.2.3; and (iii) $\Rightarrow$ (i) is 2.2.5. □

Remark 2.2.7 In 1.4.5, we saw that the change-of-base functor along F preserves (co)completeness iff F preserves pullbacks; so, combining this with 2.2.6, we see that change of base along a functor between toposes preserves (co)completeness and local smallness iff it is (what we called in A4.1.13) the inverse image of a pre-geometric morphism. This provides another sense in which (pre-)geometric morphisms (or at least their inverse image functors) are 'structure-preserving maps' between toposes: the 'structure' they preserve is the completeness and local smallness of the canonical indexing of a topos over itself. We also saw in 1.3.17 that change of base along a functor with a right adjoint preserves well-poweredness of indexed categories, and in 1.3.14 that toposes may be characterized as those cartesian categories whose canonical indexings over themselves are well-powered, so that inverse image functors really do preserve 'all the structure required to define a topos'.

We have seen in the proof of 2.2.2 that any indexed functor between locally internal categories is automatically an enriched functor. There is an interesting

special case in which this implication can be reversed: the following argument is due to R. Paré.

Proposition 2.2.8 *Let $\mathcal{S}$ be a locally cartesian closed category, and $F\colon \mathcal{S} \to \mathcal{S}$ an $\mathcal{S}$-enriched functor whose underlying ordinary functor preserves pullbacks. Then F extends to an $\mathcal{S}$-indexed functor $\mathbb{S} \to \mathbb{S}$. Moreover, this extension is unique (up to canonical isomorphism) if we require that it should preserve pullbacks as an indexed functor, and induce the given enrichment of F.*

Proof Given an object $f\colon A \to I$ of $\mathcal{S}/I$, we define $F^I(f)$ to be the morphism $u\colon P \to I$, where

$$\begin{tikzcd}
P \arrow[r, "w"] \arrow[d, "{(u,v)}"] & FA \arrow[d, "Ff"] \\
I \times F1 \arrow[r, "\tau_{I,1}"] & FI
\end{tikzcd}$$

is a pullback square (and τ is the strength of F, as in 2.1.4). Clearly, F^I is a functor $\mathcal{S}/I \to \mathcal{S}/I$. To show that the F^I form an indexed functor, let $x\colon I' \to I$ be a morphism of $\mathcal{S}$, and consider the cube

$$\begin{tikzcd}
P' \arrow[rr, "w'"] \arrow[dr] \arrow[dd, "{(u',v')}"'] & & FA' \arrow[dr] \arrow[dd, "Ff'"' near start] & \\
& P \arrow[rr, "w" near start] \arrow[dd, "{(u,v)}"' near start] & & FA \arrow[dd, "Ff"] \\
I' \times F1 \arrow[rr, "\tau_{I',1}" near start] \arrow[dr, "x \times 1"'] & & FI' \arrow[dr, "Fx"] & \\
& I \times F1 \arrow[rr, "\tau_{I,1}"] & & FI
\end{tikzcd}$$

where f' is the pullback of f along x. Since F preserves pullbacks, the right vertical face of this cube is a pullback, and the front and back faces are pullbacks by definition. So the left face is a pullback, from which it follows that $u' = F^{I'}(f')$ is (isomorphic to) the pullback of u along x. A similar calculation shows that F^I preserves pullbacks as a functor $\mathcal{S}/I \to \mathcal{S}/I$; and since $\tau_{1,1}$ is the canonical isomorphism $\pi_2\colon 1 \times F1 \to F1$, it is easy to see that F^1 is (isomorphic to) our original functor F.

Next, we must show that the enrichment of F^1 induced by the indexing is the one originally given. Let I and J be objects of $\mathcal{S}$. Then the morphism $\tilde{F}_{I,J}\colon J^I \to FJ^{FI}$ induced by the indexing is the transpose of the morphism $(J^I)^*FI \to (J^I)^*FJ$ obtained by applying $F^{(J^I)}$ to the generic morphism $(J^I)^*I \to (J^I)^*J$, i.e. the morphism

$$J^I \times I \xrightarrow{(\pi_1, \mathrm{ev})} J^I \times J .$$

Now the identification of $F^{(J^I)}(J^I)^*I$ with $(J^I)^*FI$ tells us that

$$\begin{tikzcd} J^I \times FI \arrow[r, "\tau_{J^I,I}"] \arrow[d] & F(J^I \times I) \arrow[d, "F\pi_1"] \\ J^I \times F1 \arrow[r, "\tau_{J^I,1}"] & F(J^I) \end{tikzcd}$$

is a pullback square; and similarly with I replaced by J in the second factor. So, in order to identify $\tilde{F}_{I,J}$ with $F_{I,J}$, it suffices to show that the unique morphism $J^I \times FI \to J^I \times FJ$ making the diagram

$$\begin{tikzcd} & J^I \times FI \arrow[r, "\tau_{J^I,I}"] \arrow[dl] \arrow[d, dashed] & F(J^I \times I) \arrow[d, "F(\pi_1, \mathrm{ev})"] \\ J^I \times F1 & J^I \times FJ \arrow[l] \arrow[r, "\tau_{J^I,J}"] & F(J^I \times J) \end{tikzcd}$$

commute is $(\pi_1, \overline{F_{I,J}})$, or equivalently $(\pi_1, F(\mathrm{ev}) \circ \tau_{J^I,I})$. But this is a straightforward diagram-chase, using the naturality of τ and the two commutative diagrams in the statement of 2.1.4.

For the uniqueness, we observe that for any $f\colon A \to I$ we have a pullback square

$$\begin{tikzcd} A \arrow[r, "(1,f)"] \arrow[d, "f"] & A \times I \arrow[d, "f \times 1"] \\ I \arrow[r, "\Delta"] & I \times I \end{tikzcd}$$

which we may regard as a pullback

$$\begin{tikzcd} f \arrow[r] \arrow[d] & I^*A \arrow[d, "I^*f"] \\ I^*1 \arrow[r, "\Delta"] & I^*I \end{tikzcd}$$

in $\mathcal{S}/I$. So, if we are given that F^I preserves pullbacks and that $F^I I^*$ is isomorphic to I^*F^1, the effect of F^I will be determined up to canonical isomorphism once we know what it does to the morphism $\Delta\colon I^*1 \to I^*I$. But the latter is forced upon us by the enrichment: if we think of Δ as an I-indexed family of morphisms $1 \to I$, it corresponds to a morphism $I \to I^1$ (in fact the transpose of $\pi_1\colon I \times 1 \to I$), and its image under F^I must similarly correspond to the composite of this morphism with $F_{1,I}$ (i.e. to the morphism whose transpose is $\tau_{I,1}$). □

Combining this result with 2.1.6, we obtain

Corollary 2.2.9 *If $\mathcal{S}$ is a locally cartesian closed category, then any cartesian monad on $\mathcal{S}$ extends uniquely to a cartesian $\mathcal{S}$-indexed monad on $\mathbb{S}$.* □

We do not, in fact, know any example of an enriched endofunctor of a locally cartesian closed category which fails to extend to an indexed functor. In the absence of the assumption that F preserves pullbacks, we can still prove

Lemma 2.2.10 *Let $F\colon \mathcal{S} \to \mathcal{S}$ be an $\mathcal{S}$-enriched endofunctor of a cartesian closed category $\mathcal{S}$. Then F extends, uniquely up to canonical isomorphism, to a* Sub(1)*-indexed endofunctor of the restriction of $\mathbb{S}$ to the full subcategory of $\mathcal{S}$ consisting of subterminal objects.*

Proof Let U be a subterminal object of $\mathcal{S}$. Since the composite

$$\mathcal{S}/U \xrightarrow{\Sigma_U} \mathcal{S} \xrightarrow{U^*} \mathcal{S}/U$$

is isomorphic to the identity, it is easy to see that if we have any $F^U : \mathcal{S}/U \to \mathcal{S}/U$ satisfying $F^U U^* \cong U^* F$, then $F^U \cong U^* F \Sigma_U$. Taking the latter as a definition of F^U, the verification that it yields a Sub(1)-indexed functor amounts to showing that for any object I of $\mathcal{S}$ we have $U \times F(I) \cong U \times F(U \times I)$. But the morphism $(\pi_1, \tau_{U,I})$, where τ is the strength of F as in 2.1.4, is readily checked to be a two-sided inverse for $1 \times F(\pi_2)$. □

Corollary 2.2.11 *If $\mathcal{S}$ is a cocomplete, locally small topos in which the subterminal objects form a class of generators, then any $\mathcal{S}$-enriched functor $F\colon \mathcal{S} \to \mathcal{S}$ extends, uniquely up to isomorphism, to an $\mathcal{S}$-indexed endofunctor.*

Proof Let I be an object of $\mathcal{S}$. By the assumptions on $\mathcal{S}$, I can be expressed as a colimit of a small diagram $D\colon \mathcal{J} \to \mathcal{S}$, whose vertices are subterminal objects. Given any object $(A \to I)$ of $\mathcal{S}/I$, A is similarly a colimit of the diagram $j \mapsto D(j) \times_I A$, since colimits in $\mathcal{S}$ are stable under pullback. Applying $F^{D(j)}$ (defined as in 2.2.10) to the jth vertex of this diagram, and using the fact that the $F^{D(j)}$ form a Sub(1)-indexed functor, we obtain another diagram of type $\mathcal{J}$ in $\mathcal{S}$, such that for any morphism $j \to j'$ in $\mathcal{J}$ we have a commutative square (in fact a pullback)

$$\begin{tikzcd}
F^{D(j)}(D(j) \times_I A) \arrow[r] \arrow[d] & F^{D(j')}(D(j') \times_I A) \arrow[d] \\
D(j) \arrow[r] & D(j')
\end{tikzcd} .$$

Taking the colimit of this diagram, we obtain an object of $\mathcal{S}/I$, which we define to be $F^I(A \to I)$. The verification that this construction defines an $\mathcal{S}$-indexed functor, and that it extends the Sub(1)-indexed functor already defined in 2.2.10, is straightforward. The uniqueness of the extension again follows from the stability of colimits under pullback: $F^I(A \to I)$ must be the colimit of its pullbacks along the morphisms $D(j) \to I$. □

The hypotheses on $\mathcal{S}$ in 2.2.11 are equivalent to saying that it admits a localic geometric morphism to **Set**; cf. A4.1.9 and C1.4.7.

Suggestions for further reading: Betti [108, 114], Penon [955].

B2.3 Internal categories and diagram categories

As we emphasized in Section A1.1, the theory of categories is an elementary one, and so may be interpreted in the internal logic of any category $\mathcal{S}$ with sufficient structure. In fact, since the theory is a cartesian one, as defined in D1.3.4, all we need is that $\mathcal{S}$ should have finite limits. Throughout this section $\mathcal{S}$ will be (at least) a cartesian category.

Definition 2.3.1 (*a*) An *internal category* $\mathbb{C}$ in $\mathcal{S}$ is specified by the following data:

(i) three objects C_0 (the *object of objects* of $\mathbb{C}$), C_1 (the *object of morphisms*) and C_2 (the *object of composable pairs*) of $\mathcal{S}$;

(ii) two morphisms $d_0^1, d_1^1\colon C_1 \to C_0$ (interpreted as *codomain* and *domain* respectively), a morphism $s_0^0\colon C_0 \to C_1$ (*inclusion of identities*) and three morphisms $d_0^2, d_1^2, d_2^2\colon C_2 \to C_1$ (respectively the *first member*, the *composite* and the *second member* of the composable pair);

subject to the conditions that

$$\begin{tikzcd} C_2 \arrow[r, "d_2^2"] \arrow[d, "d_0^2"'] & C_1 \arrow[d, "d_0^1"] \\ C_1 \arrow[r, "d_1^1"] & C_0 \end{tikzcd}$$

is a pullback, that $d_0^1 s_0^0 = 1_{C_0} = d_1^1 s_0^0$, that $d_0^1 d_0^2 = d_0^1 d_1^2$ and $d_1^1 d_1^2 = d_1^1 d_2^2$, that $d_1^2 s_0^1 = 1_{C_1} = d_1^2 s_1^1$ (where $s_0^1 \colon C_1 \to C_2$ is the morphism into the pullback induced by $s_0^0 d_0^1 \colon C_1 \to C_1$ and 1_{C_1}, and s_1^1 is similarly defined), and that d_1^2 is *associative* in the sense that, if we form the *object of composable triples* C_3 as the limit of the diagram

$$\begin{tikzcd} C_1 \arrow[dr, "d_1^1"'] & & C_1 \arrow[dl, "d_0^1"] \arrow[dr, "d_1^1"'] & & C_1 \arrow[dl, "d_0^1"] \\ & C_0 & & C_0 & \end{tikzcd}$$

then two possible composites $C_3 \to C_2 \to C_1$ are equal.

(*b*) If $\mathbb{C}$ and $\mathbb{D}$ are internal categories, an *internal functor* $f \colon \mathbb{C} \to \mathbb{D}$ consists of morphisms $f_0 \colon C_0 \to D_0$, $f_1 \colon C_1 \to D_1$ and $f_2 \colon C_2 \to D_2$ commuting with the structure morphisms in (*a*)(ii) above.

(c) If f and $g \colon \mathbb{C} \to \mathbb{D}$ are internal functors, an *internal natural transformation* $\alpha \colon f \to g$ consists of a morphism $\alpha \colon C_0 \to D_1$ such that $d_0^1 \alpha = g_0$, $d_1^1 \alpha = f_0$, and

$$d_1^2(\alpha d_0^1, f_1) = d_1^2(g_1, \alpha d_1^1) \colon C_1 \longrightarrow D_1 \ .$$

The internal categories, internal functors and internal natural transformations in $\mathcal{S}$ form a 2-category, which we denote $\mathfrak{Cat}(\mathcal{S})$. We also write $\mathbf{Cat}(\mathcal{S})$ for the underlying 1-category of $\mathfrak{Cat}(\mathcal{S})$.

The reader may readily verify that $\mathfrak{Cat}(\mathbf{Set})$ is just the 2-category $\mathfrak{Cat}$ of small categories, as previously defined. We also note that a cartesian functor $\mathcal{S} \to \mathcal{T}$ induces a 2-functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Cat}(\mathcal{T})$, in an obvious way. Moreover, $\mathbf{Cat}(\mathcal{S})$ is itself a cartesian category; in fact the forgetful functor $\mathbf{Cat}(\mathcal{S}) \to \mathcal{S}^3$ which sends $\mathbb{C}$ to (C_0, C_1, C_2) creates finite limits.

Remark 2.3.2 There is some redundancy in Definition 2.3.1, in that the object C_2 and the morphisms d_0^2 and d_2^2 do not need to be (and in general, will not be) explicitly specified as part of the definition of an internal category, being recoverable from the rest via the pullback diagram contained in the definition. Similarly,

when we define an internal functor we do not need to specify f_2 explicitly, since it is uniquely determined by f_0 and f_1.

Nevertheless, the notation we used was intended to be suggestive that the definition is in a sense incomplete. If we define C_n, the *object of composable n-tuples* of morphisms in $\mathbb{C}$, by an appropriate finite limit diagram in $\mathcal{S}$ (generalizing that which appeared in the definition for the case $n = 3$), then we have $(n+1)$ morphisms $d_i^n \colon C_n \to C_{n-1}$ $(0 \leq i \leq n)$, of which the first and last represent the operations of dropping the last and first members of a composable n-tuple, and the others represent those of composing one of the $(n-1)$ adjacent pairs. Similarly, we have n morphisms $s_j^{n-1} \colon C_{n-1} \to C_n$, which represent the insertion of identity morphisms at each possible point in an $(n-1)$-tuple. Moreover, these morphisms satisfy the *simplicial identities*

$$\begin{aligned}
d_i^n d_{j+1}^{n+1} &= d_j^n d_i^{n+1} \quad (i \leq j)\ , \\
s_{j+1}^n s_i^{n-1} &= s_i^n s_j^{n-1} \quad (i \leq j)\ , \text{ and} \\
d_i^{n+1} s_j^n &= s_{j-1}^{n-1} d_i^n \quad (i < j) \\
&= 1_{C_n} \quad (j \leq i \leq j+1) \\
&= s_j^{n-1} d_{i-1}^n \quad (i > j+1);
\end{aligned}$$

that is, they make $(n \mapsto C_n)$ into a *simplicial object* in $\mathcal{S}$, i.e. a functor $\mathbf{\Delta}^{\mathrm{op}} \to \mathcal{S}$, where $\mathbf{\Delta}$ is the category of nonempty finite totally ordered sets and order-preserving maps. (Note: as an object of $\mathbf{\Delta}$, n denotes the $(n+1)$-element totally ordered set $\{0, 1, \ldots, n\}$; this 'dimension shift' is not only traditional but seemingly unavoidable.) Similarly, an internal functor $f \colon \mathbb{C} \to \mathbb{D}$ extends uniquely to a *morphism of simplicial objects*, i.e. a natural transformation between the corresponding functors $\mathbf{\Delta}^{\mathrm{op}} \to \mathcal{S}$; thus we can identify the 1-category $\mathbf{Cat}(\mathcal{S})$ with a full subcategory of the category $\mathbf{Simpl}(\mathcal{S}) = [\mathbf{\Delta}^{\mathrm{op}}, \mathcal{S}]$ of simplicial objects in $\mathcal{S}$. (Although we shall not need this fact, it is worth noting that the objects in this subcategory are, up to isomorphism, precisely the functors $\mathbf{\Delta}^{\mathrm{op}} \to \mathcal{S}$ which preserve those pullbacks which exist in $\mathbf{\Delta}^{\mathrm{op}}$.) The notion of internal natural transformation also corresponds to a simplicial concept, namely that of a simplicial homotopy between morphisms of simplicial objects; however, simplicial homotopies cannot be composed in general, so $\mathbf{Simpl}(\mathcal{S})$ cannot be extended to a 2-category.

Having got the formal definition out of the way, we shall henceforth omit the superscripts from the names of the morphisms d_i^n and s_j^n whenever possible; the context will usually make it clear which 'dimension' is intended. The simplicial object corresponding to an internal category is commonly called its *nerve*, but we shall not find it necessary to distinguish notationally between the two; for us, an internal category may be identified with its nerve whenever it is profitable to do so.

Next, we tackle the relationship between internal and indexed/locally internal categories. If $\mathbb{C}$ is an internal category in $\mathcal{S}$ and I is an object of $\mathcal{S}$, we write C^I for the ordinary category whose objects are the morphisms $I \to C_0$ in $\mathcal{S}$, and whose morphisms are the morphisms $I \to C_1$, the operations 'domain', 'codomain', 'identity morphism' and 'composite' being induced in the obvious way by composition with the structural morphisms of $\mathbb{C}$. If $x: I \to J$ in $\mathcal{S}$, then composition (on the other side) with x induces a functor $x^*: C^J \to C^I$; the assignment $x \mapsto x^*$ is (strictly) functorial, and so $I \mapsto C^I$ becomes an $\mathcal{S}$-indexed category. Once again, we shall usually not bother to distinguish notationally between this indexed category and the internal category $\mathbb{C}$; but when it is necessary to make the distinction we shall denote the indexed category by $[\mathbb{C}]$. The justification for this abuse of notation is contained in

Lemma 2.3.3 *The assignment* $\mathbb{C} \mapsto [\mathbb{C}]$ *is a full embedding of 2-categories* $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{CAT}_{\mathcal{S}}$.

Proof It is clear that an internal functor $f: \mathbb{C} \to \mathbb{D}$ induces, by composition, an indexed functor $[\mathbb{C}] \to [\mathbb{D}]$; and similarly for internal natural transformations. So the assignment is 2-functorial. The fact that, for each pair $(\mathbb{C}, \mathbb{D})$, it induces an equivalence of categories between $\mathfrak{Cat}(\mathcal{S})(\mathbb{C}, \mathbb{D})$ and $[[\mathbb{C}], [\mathbb{D}]]$ is an easy exercise in the use of the Yoneda lemma: for example, given an indexed functor $F: [\mathbb{C}] \to [\mathbb{D}]$, we define $f_0 = F^{C_0}(1_{C_0}): C_0 \to D_0$ and $f_1 = F^{C_1}(1_{C_1}): C_1 \to D_1$, and verify that these define an internal functor $\mathbb{C} \to \mathbb{D}$; then the coherence isomorphisms for F yield an indexed natural isomorphism between F and the indexed functor induced by f. □

The reader should be warned not to read into Lemma 2.3.3 more than it actually says: the functor is a full embedding of 2-categories, but not a full embedding of 1-categories. In particular, although it follows that any indexed equivalence between $[\mathbb{C}]$ and $[\mathbb{D}]$ must be induced by an internal equivalence between $\mathbb{C}$ and $\mathbb{D}$, it is possible for $[\mathbb{C}]$ and $[\mathbb{D}]$ to be isomorphic when $\mathbb{C}$ and $\mathbb{D}$ are not. The following simple example is due to R. Paré.

Example 2.3.4 Let C and D be objects of a cartesian category $\mathcal{S}$ which 'provide a counterexample to the Cantor–Bernstein theorem in $\mathcal{S}$': that is, such that there are monomorphisms $u: C \rightarrowtail D$ and $v: D \rightarrowtail C$ but no isomorphism $C \cong D$. (Such pairs of objects can easily be found in suitable non-Boolean toposes; see D4.1.12.) Let $\mathbb{C}$ be the *indiscrete* internal category with object of objects C (that is, take $C_n = C^{n+1}$ for all n, with the d_i taken to be appropriate product projections and the s_j induced by the diagonal map), and let $\mathbb{D}$ be similarly constructed from D. Then it is clear that $\mathbb{C}$ and $\mathbb{D}$ are equivalent, but not isomorphic, in $\mathfrak{Cat}(\mathcal{S})$. However, for each I, composition with u and v induces injections in either direction between $\mathcal{S}(I, C)$ and $\mathcal{S}(I, D)$, so we can use the (external) Cantor–Bernstein theorem to find a bijection between these sets, which will be an isomorphism between the (indiscrete) categories C^I and $\mathcal{D}^I$. We cannot hope to choose these bijections so that they are exactly compatible for all

I (that is, commute with composition with a fixed $x: I \to J$); but, again because the categories are indiscrete, they will do so up to unique natural isomorphism, and so we have defined not just an equivalence but an isomorphism of indexed categories between $[\mathbb{C}]$ and $[\mathbb{D}]$.

The moral of this example, if it has one, is that isomorphism of indexed categories is not really a sensible concept to study: the definition of indexed category is so 'loose' that $\mathfrak{CAT}_{\mathcal{S}}$ has to be considered as a 2-category, not as a 1-category. We shall say that an $\mathcal{S}$-indexed category is *essentially small* if it is equivalent to $[\mathbb{C}]$ for some internal category $\mathbb{C}$; this replaces the unsatisfactory notion of essential smallness which we discussed in 1.3.13.

It is easy to see that an indexed category of the form $[\mathbb{C}]$ is locally small: if $x: I \to C_0$ and $y: J \to C_0$ are two objects of it, then the object of $\mathcal{S}/I \times J$ which indexes the morphisms from x to y in $[\mathbb{C}]$ is the left vertical arrow in the pullback

$$\begin{tikzcd} P \arrow[r] \arrow[d] & C_1 \arrow[d, "{(d_1, d_0)}"] \\ I \times J \arrow[r, "x \times y"] & C_0 \times C_0 \end{tikzcd}$$

(and the generic morphism itself is the top edge of this square). In the opposite direction, we have the following important method of constructing internal categories from locally small indexed categories, which helps to explain why the latter are also called locally internal categories.

Definition 2.3.5 Let $\mathbb{C}$ be a locally small $\mathcal{S}$-indexed category, and let A be an object of C^I for some I. We define the *internal full subcategory* of $\mathbb{C}$ generated by A, which we denote by $\mathbb{C}[A]$, as follows. Its object of objects $C_0[A]$ is simply I, and $(d_1, d_0): C_1[A] \to C_0[A] \times C_0[A]$ is the object of $\mathcal{S}/I \times I$ indexing morphisms from A to A. The morphism $s_0: C_0[A] \to C_1[A]$ corresponds (modulo coherence isomorphisms) to the identity morphism $A \to A$ in C^I, and $d_1: C_2[A] \to C_1[A]$ similarly corresponds to the composite of the morphisms in $C^{C_2[A]}$ obtained by pulling back the generic morphism along d_0 and d_2. It is straightforward to verify that the foregoing data does define an internal category.

$\mathbb{C}[A]$ comes equipped with a canonical indexed functor (from the indexed category corresponding to it) to the original category $\mathbb{C}$: this sends an object $x: J \to I$ to $x^*(A) \in \text{ob } C^J$, and a morphism $y: J \to C_1[A]$ to the morphism $(d_1 y)^*(A) \to (d_0 y)^*(A)$ in C^J which corresponds to it. It is easy to see that this indexed functor is full and faithful; hence the name 'internal full subcategory'.

We recall that, given a functor $F: \mathcal{S} \to \mathcal{T}$, we have a 're-indexing' functor F^* from $\mathcal{T}$-indexed categories to $\mathcal{S}$-indexed categories; on the other hand, given an

internal category $\mathbb{C}$ in $\mathcal{T}$, we require a (pullback-preserving) functor $G: \mathcal{T} \to \mathcal{S}$ to produce an internal category $G(\mathbb{C})$ in $\mathcal{S}$. The two constructions are linked in an obvious way:

Lemma 2.3.6 *Let $F: \mathcal{S} \to \mathcal{T}$ be a functor having a right adjoint G, and let $\mathbb{C}$ be an internal category in $\mathcal{T}$. Then the $\mathcal{S}$-indexed categories $F^*[\mathbb{C}]$ and $[G(\mathbb{C})]$ are equivalent.*

Proof For any object I of $\mathcal{S}$, we have a bijection between morphisms $F(I) \to C_i$ ($i = 0$ or 1) in $\mathcal{T}$ and morphisms $I \to G(C_i)$ in $\mathcal{S}$. This sets up an equivalence (which is in fact an isomorphism) between the two $\mathcal{S}$-indexed categories. □

We shall most frequently use this result when $\mathcal{S}$ and $\mathcal{T}$ are toposes, and F and G are the inverse and direct images of a geometric morphism. We shall also need

Corollary 2.3.7 *Let $G: \mathcal{S} \to \mathcal{T}$ be a functor having a left adjoint F which preserves pullbacks (for example, the direct image of a pre-geometric morphism, as defined in* A4.1.13*). If $\mathbb{C}$ is an internal category in $\mathcal{S}$ which is $\mathcal{S}$-complete (resp. $\mathcal{S}$-cocomplete) as an indexed category, then $G(\mathbb{C})$ is $\mathcal{T}$-complete (resp. $\mathcal{T}$-cocomplete).*

Proof This follows from 2.3.6 and 1.4.9. □

There is a well-known result due to P. Freyd (cf. [487]) which says that a complete small category must be a preorder. The proof of this result requires the law of excluded middle, and so one might expect it to fail for indexed categories over a non-Boolean base topos $\mathcal{S}$; we shall see that it does indeed fail in certain cases, in Chapter F2. However, in the main we shall be interested in applying 2.3.7 to *internal posets*, that is to internal categories $\mathbb{C}$ such that $(d_0, d_1): C_1 \to C_0 \times C_0$ is monic and

$$\begin{tikzcd} C_0 \arrow[r, "s_0"] \arrow[d, "s_0"'] & C_1 \arrow[d, "{(d_0, d_1)}"] \\ C_1 \arrow[r, "{(d_1, d_0)}"] & C_0 \times C_0 \end{tikzcd}$$

is a pullback.

Examples 2.3.8 (*a*) One particular internal poset, in any topos, with which we are already familiar is the subobject classifier Ω, equipped with the partial ordering Ω_1 defined in A1.6.3. Of course, the indexed category corresponding to this internal category is (up to equivalence) the category $\mathbb{S}\mathsf{ub}$ of 1.2.2(g), and we observed before 1.4.7 that $\mathbb{S}\mathsf{ub}$ is complete and cocomplete over any Heyting category (in particular, over any topos); so we deduce that Ω is a complete and cocomplete internal poset in any topos. Applying 2.3.7 with $G = (-)^A$ and $F = (-) \times A$, we deduce that any power object PA in a topos, made into a poset

$\mathbb{P}A$ with its canonical ordering defined in A2.3.2, is complete and cocomplete; applying it to the direct and inverse images of a geometric morphism $f: \mathcal{S} \to \mathcal{T}$, we deduce that $f_*(\Omega)$ (and more generally $f_*(\mathbb{P}A)$ for any object A of $\mathcal{S}$) is a complete and cocomplete internal poset in $\mathcal{T}$.

(*b*) Another example of interest derives from the notion of local operator which we studied in Sections A4.4 and A4.5. The external poset $\mathbf{Lop}(\mathcal{S})$ which we defined in A4.5.10 may be made into an indexed poset $\mathbb{L}\mathrm{op}(\mathcal{S})$, by setting $\mathbf{Lop}(\mathcal{S})^I = \mathbf{Lop}(\mathcal{S}/I)$ (and, for a local operator $j: I^*\Omega \to I^*\Omega$ on $\mathcal{S}/I$, $x^*(j)$ is simply the pullback of j along $x: K \to I$ in $\mathcal{S}$). It is easy to see that we may construct a subobject $L_0 \rightarrowtail \Omega^\Omega$, as an intersection of three appropriate equalizers, such that a morphism $I \to \Omega^\Omega$ factors through L_0 iff its transpose $I^*\Omega \to I^*\Omega$ is a local operator on $\mathcal{S}/I$; hence we can construct an internal poset $\mathbb{L}$ such that $[\mathbb{L}] \simeq \mathbb{L}\mathrm{op}(\mathcal{S})$. Moreover, the construction of the local operator generated by a subobject of Ω, given in A4.5.13(i), commutes with pullback functors, so it defines an $\mathcal{S}$-indexed (and hence internal) left adjoint for the inclusion $\mathbb{L} \to \mathbb{P}\Omega$, from which it follows that $\mathbb{L}$ is $\mathcal{S}$-complete and cocomplete (cf. 2.3.9 below). We note in passing that the constructions of A4.5.19 also give rise to $\mathcal{S}$-indexed adjunctions between $\mathbb{L}$ and Ω.

For internal posets in a topos, we have an alternative characterization of (co)completeness in terms of 'taking sups of subsets' rather than 'taking sups of indexed families'. Given a poset $\mathbb{A} = (A_1 \rightrightarrows A)$ in a topos $\mathcal{S}$, let $\downarrow: A \to PA$ be the name (in the sense of A2.1.1) of the relation $(d_0, d_1): A_1 \rightarrowtail A \times A$ (note that this is the *opposite* of the order-relation as we usually think of it). If $\mathcal{S} =$ **Set**, $\downarrow$ is the mapping which sends a to the principal ideal $\{b \in A \mid b \leq a\}$, which explains the notation. It is easy to verify that $\downarrow$ is an order-preserving map $\mathbb{A} \to \mathbb{P}A$.

Lemma 2.3.9 *For an internal poset $\mathbb{A}$ in a topos $\mathcal{S}$, the following are equivalent:*

(i) *$\mathbb{A}$ is cocomplete, i.e. $[\mathbb{A}]$ is an $\mathcal{S}$-cocomplete indexed category.*

(ii) *$\downarrow: \mathbb{A} \to \mathbb{P}A$ has a left adjoint $\bigvee: \mathbb{P}A \to \mathbb{A}$ in the 2-category of internal posets in $\mathcal{S}$.*

(iii) *$\downarrow$ has an order-preserving left inverse.*

(iv) *$\mathbb{A}$ is a retract of a poset of the form $\mathbb{P}B$.*

Proof (i) $\Rightarrow$ (ii): We define $\bigvee: PA \to A$ to be $\Sigma_n(e)$, where $(n, e): \in_A \rightarrowtail PA \times A$ is as usual the generic relation $PA \nrightarrow A$. We must first verify that $\bigvee$ is order-preserving. Suppose given maps $x_1, x_2: I \rightrightarrows PA$ satisfying

$x_1 \leq x_2$; this means that if we form the pullbacks

$$\begin{tikzcd} J_i \arrow[r, "q_i"] \arrow[d, "p_i"'] & \in_A \arrow[d, "n"] \\ I \arrow[r, "x_i"] & PA \end{tikzcd}$$

($i = 1, 2$), then $J_1 \leq J_2$ as subobjects of $I \times A$; that is, we have a monomorphism $m\colon J_1 \rightarrowtail J_2$ satisfying $p_2 m = p_1$ and $eq_2 m = eq_1$. Now we have $\bigvee x_i = x_i^* \Sigma_n(e) = \Sigma_{p_i} q_i^*(e)$ by the Beck–Chevalley condition; but

$$\Sigma_{p_1} q_1^*(e) = \Sigma_{p_2} \Sigma_m m^* q_2^*(e) \leq \Sigma_{p_2} q_2^*(e)$$

since $(\Sigma_m \dashv m^*)$ and Σ_{p_2} is order-preserving.

Now we verify that $\bigvee$ is left adjoint to $\downarrow$. We have a pullback diagram

$$\begin{tikzcd} A_1 \arrow[r, "z"] \arrow[d, "d_0"'] & \in_A \arrow[d, "n"] \\ A \arrow[r, "\downarrow"] & PA \end{tikzcd}$$

by the definition of $\downarrow$, where the composite ez is d_1; so by applying the Beck–Chevalley condition to this square we obtain $\bigvee \downarrow = \Sigma_{d_0} d_1$. But since $d_1 \leq d_0$ in the poset $\mathcal{S}(A_1, A)$, we have $\Sigma_{d_0} d_1 \leq \Sigma_{d_0} d_0 = \Sigma_{d_0} d_0^*(1_A) \leq 1_A$. On the other hand, we have $\bigvee n = n^* \Sigma_n(e) \geq e$; that is, the morphism $(\bigvee n, e)\colon \in_A \to A \times A$ factors through (d_0, d_1) (say by $y\colon \in_A \to A_1$). Thus we have a commutative diagram

$$\begin{tikzcd} \in_A \arrow[r, "y"] \arrow[d, "n"'] & A_1 \arrow[r, "z"] \arrow[d, "d_0"] & \in_A \arrow[r, "e"] \arrow[d, "n"] & A \\ PA \arrow[r, "\bigvee"] & A \arrow[r, "\downarrow"] & PA & \end{tikzcd}$$

where the top composite is e, so we conclude that (n, e) is less than or equal to the relation named by $\downarrow \bigvee$ in the poset of relations $PA \nrightarrow A$, i.e. $1_{PA} \leq \downarrow \bigvee$ in the poset of maps $PA \to PA$. Thus we have constructed the unit and counit of the required adjunction; the triangular identities are trivial, since we are dealing with posets.

(ii) $\Rightarrow$ (iii): It follows easily from reflexivity and antisymmetry of the order-relation that $\downarrow$ is monic; so the counit $\bigvee \downarrow \leq 1_A$ of the adjunction is necessarily an equality. And (iii) $\Rightarrow$ (iv) is trivial.

(iv) ⇒ (i): A poset is trivially Cauchy-complete, since it has no nontrivial idempotents; so 1.1.10 enables us to lift the left adjoints for the transition maps of $[\mathbb{P}B]$ to left adjoints for those of $[\mathbb{A}]$. The Beck–Chevalley condition for $[\mathbb{A}]$ also follows trivially from that for $[\mathbb{P}B]$; so $[\mathbb{A}]$ has $\mathcal{S}$-indexed coproducts. The fact that the individual posets $[A]^I = \mathcal{S}(I, A)$ have finite coproducts (joins) may be deduced from 1.4.3 (or from the fact that the posets $[PB]^I$ have them); and equalizers are again trivial because we are dealing with posets. □

Note that the proof of (iv) ⇒ (i) in 2.3.9 shows that a cocomplete internal poset in a topos is necessarily also complete, since $[\mathbb{P}B]$ is both complete and cocomplete. We shall prove the corresponding result for arbitrary internal categories in the next section. It is also possible to give a direct proof of (ii) ⇒ (i): given $\bigvee$, and a pair of morphisms $x\colon J \to I$ and $u\colon J \to A$ in $\mathcal{S}$, we define $\Sigma_x(u)$ to be the composite

$$I \xrightarrow{\ulcorner x^{\bullet} \urcorner} PJ \xrightarrow{\exists u} PA \xrightarrow{\bigvee} A\,,$$

where $x^{\bullet}$, as in A3.1.3, denotes the relation $I \nrightarrow J$ which is tabulated by $(x, 1_J)\colon J \rightarrowtail I \times J$. We omit the detailed verification that this has the required properties.

For future reference, we note a couple of applications of 2.3.9:

Corollary 2.3.10

(i) *A logical functor $F\colon \mathcal{S} \to \mathcal{T}$ between toposes preserves (co)completeness of internal posets.*

(ii) *If j is a local operator on a topos $\mathcal{S}$, then an internal poset in $\mathcal{S}$ is* $\mathbf{sh}_j(\mathcal{S})$*-cocomplete iff it is $\mathcal{S}$-cocomplete and its underlying object is a j-sheaf.*

Proof (i) Since F preserves power objects, it clearly preserves any of the conditions (ii), (iii) and (iv) of 2.3.9.

(ii) One direction is immediate from 2.3.7, since the inclusion $\mathbf{sh}_j(\mathcal{S}) \to \mathcal{S}$ is a direct image functor. Conversely, suppose that $\mathbb{A}$ is cocomplete as an internal poset in $\mathcal{S}$, and that its underlying object A is a j-sheaf. Then its order-relation A_1 is also a j-sheaf (since it may be expressed as the equalizer of

$$A \times A \underset{\vee}{\overset{\pi_2}{\rightrightarrows}} A$$

where $\vee$ is the binary join map) and hence j-closed in $A \times A$ by A4.3.8; so the classifying map of $A_1 \rightarrowtail A \times A$ factors through $\Omega_j \rightarrowtail \Omega$, and hence $\downarrow$ factors through the subobject $P_j A = (\Omega_j)^A \rightarrowtail PA$ which is the power object of A in $\mathbf{sh}_j(\mathcal{S})$. So this factorization, and the restriction of $\bigvee\colon PA \to A$ to $P_j A$, express $\mathbb{A}$ as a retract of a power object $\mathbb{P}_j A$ in the category of internal posets in $\mathbf{sh}_j(\mathcal{S})$. □

Next, we turn to the study of indexed functors from internal categories to arbitrary indexed categories. Once again, the Yoneda lemma enables us to describe these in simpler terms:

Definition 2.3.11 Let $\mathbb{C}$ be an internal category in a cartesian category $\mathcal{S}$, and $\mathbb{D}$ an indexed category over $\mathcal{S}$. By a *diagram* of shape $\mathbb{C}$ in $\mathbb{D}$, we mean a pair (F, ϕ), where F is an object of $\mathcal{D}^{C_0}$ and $\phi\colon d_1^*F \to d_0^*F$ is a morphism in $\mathcal{D}^{C_1}$, such that $s_0^*\phi$ is (modulo coherence isomorphisms) the identity morphism $F \to F$, and $d_1^*\phi$ is similarly the composite of $d_0^*\phi$ and $d_2^*\phi$ in $\mathcal{D}^{C_2}$. A *morphism of diagrams* $(F, \phi) \to (G, \psi)$ is a morphism $F \to G$ in $\mathcal{D}^{C_0}$ which commutes in the obvious sense with ϕ and ψ. We shall write $\mathcal{D}^{\mathbb{C}}$ (or $[\mathbb{C}, \mathcal{D}]$, according to context) for the category of diagrams of shape $\mathbb{C}$ in $\mathbb{D}$ and morphisms between them.

Examples 2.3.12 (*a*) For any object I of $\mathcal{S}$, the *discrete* internal category $\mathbb{I}$ corresponding to I has its objects of objects and morphisms both equal to I and all its structure morphisms equal to 1_I; it is easy to see that the corresponding indexed category is the one which we called $\mathbb{I}$ in 1.2.5. Moreover, if (F, ϕ) is a diagram of shape $\mathbb{I}$, then it is easy to see that ϕ must be the identity morphism on F; so the diagram category $\mathcal{D}^{\mathbb{I}}$ is isomorphic to $\mathcal{D}^I$. In fact we can regard the assignment $I \mapsto \mathbb{I}$ as a (full and faithful) functor $\mathcal{S} \to \mathbf{Cat}(\mathcal{S})$; and, since an internal functor $f\colon \mathbb{C} \to \mathbb{C}'$ induces a functor $f^*\colon \mathcal{D}^{\mathbb{C}'} \to \mathcal{D}^{\mathbb{C}}$ in an obvious way, we see that the assignment $\mathbb{C} \mapsto \mathcal{D}^{\mathbb{C}}$ yields a canonical extension of an arbitrary $\mathcal{S}$-indexed category $\mathbb{D}$ to a $\mathbf{Cat}(\mathcal{S})$-indexed one. Similarly, any $\mathcal{S}$-indexed functor can be canonically extended to a $\mathbf{Cat}(\mathcal{S})$-indexed functor.

From now on, we shall tend to identify objects of $\mathcal{S}$ with discrete internal categories whenever we need to do so, without explicitly commenting on the fact. For example, we make $\mathcal{D}^{\mathbb{C}}$ into an $\mathcal{S}$-indexed category $\mathbb{D}^{\mathbb{C}}$ by setting

$$(\mathcal{D}^{\mathbb{C}})^I = \mathcal{D}^{I \times \mathbb{C}}$$

(where the exponent on the right really means the product of the given category $\mathbb{C}$ and the discrete category $\mathbb{I}$). Equivalently, $(\mathcal{D}^{\mathbb{C}})^I$ may be identified with the category of diagrams of shape $\mathbb{C}$ in the indexed category $\mathbb{D}^I$, as defined in 1.2.2(*f*).

(*b*) If $\mathcal{S}$ is a positive coherent category, then any finite (external) category $\mathcal{C}$ may be identified with an internal category in $\mathcal{S}$, via the coherent functor $\Delta\colon \mathbf{Set}_f \to \mathcal{S}$ of A1.4.7. If $\mathbb{D}$ is an $\mathcal{S}$-indexed category which is a stack for the coherent coverage, so that $\mathcal{D}^{\Delta I}$ is an I-indexed product of copies of $\mathcal{D}^1$ for each finite set I, then it is easy to see that a diagram of shape $\Delta\mathcal{C}$ in $\mathbb{D}$ is essentially the same thing as an ordinary functor $\mathcal{C} \to \mathcal{D}^1$. More generally, if $\mathcal{S}$ is an ∞-positive geometric category, and $\mathbb{D}$ is a stack for the coverage derived from arbitrary coproducts in $\mathcal{S}$, we can similarly identify functors $\mathcal{C} \to \mathcal{D}^1$ with diagrams whose shape is an internal category in $\mathcal{S}$, for any small category $\mathcal{C}$.

(c) Of course, we say that an internal category $\mathbb{C}$ is a *groupoid* if there exists a morphism $t\colon C_1 \to C_1$ satisfying commutative diagrams which correspond to the

assertion that t sends each morphism of $\mathbb{C}$ to its inverse. For future reference, we note that if $\mathbb{C}$ is a groupoid, then in addition to the pullback square appearing in Definition 2.3.1(a), we also have pullback squares

$$\begin{tikzcd} C_2 \arrow[r, "d_1"] \arrow[d, "d_0"] & C_1 \arrow[d, "d_0"] \\ C_1 \arrow[r, "d_0"] & C_0 \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} C_2 \arrow[r, "d_2"] \arrow[d, "d_1"] & C_1 \arrow[d, "d_1"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd} \quad ;$$

for if $f, g\colon I \rightrightarrows C_1$ are morphisms satisfying $d_0 f = d_0 g$, then it is easily verified that the factorization of $(f, d_1(tf, g))\colon I \to C_1 \times C_1$ through C_2 is the unique factorization of (f, g) through (d_0, d_1), and similarly for the right-hand square. Note also that an equivalence relation on an object of $\mathcal{S}$, as defined in A1.3.6, is an internal groupoid: indeed, equivalence relations are exactly those groupoids $\mathbb{G}$ which are also preorders (i.e. such that $(d_0, d_1)\colon G_1 \to G_0 \times G_0$ is monic). If $\mathbb{G}$ is an effective equivalence relation (i.e. occurs as the kernel-pair of some morphism $x\colon G_0 \to I$), then the notion of a diagram of shape $\mathbb{G}$ in $\mathbb{D}$ is exactly that of an object of $\mathcal{D}^{G_0}$ equipped with descent data for the singleton family (x), as we defined it in Section B1.5.

Lemma 2.3.13 *Let $\mathbb{C}$ be an internal category in a cartesian category $\mathcal{S}$, and $\mathbb{D}$ an indexed category over $\mathcal{S}$. Then $\mathcal{D}^{\mathbb{C}}$ is equivalent to the category $[\mathbb{C}, \mathbb{D}]$ of indexed functors $\mathbb{C} \to \mathbb{D}$ and indexed natural transformations between them, as defined in Section* B1.2.

Proof Given an indexed functor $G\colon \mathbb{C} \to \mathbb{D}$, we define $F = G^{C_0}(1_{C_0})$ and $\phi = G^{C_1}(1_{C_1})$. Conversely, given a diagram (F, ϕ), we define G by setting $G^I(x\colon I \to C_0) = x^*F$ and $G^J(y\colon J \to C_1) = y^*\phi\colon (d_1 y)^*F \to (d_0 y)^*F$. The remaining details are straightforward. □

We note that Example 1.2.5 was a special case of the above lemma, in which the internal category $\mathbb{C}$ was taken to be discrete, as in 2.3.12(a). The lemma may be extended in an obvious way to an equivalence between $\mathbb{D}^{\mathbb{C}}$ and the indexed functor category $[\![\mathbb{C}, \mathbb{D}]\!]$ of Section B1.2.

A corollary of 2.3.13, which will be of use in Section B3.2, is the following:

Corollary 2.3.14 *Let $G\colon \mathcal{T} \to \mathcal{S}$ be a functor preserving pullbacks; let $\mathbb{C}$ be an internal category in $\mathcal{T}$, and $\mathbb{D}$ a $\mathcal{S}$-indexed category. Then the category $[G(\mathbb{C}), \mathbb{D}]$ is equivalent to $[\mathbb{C}, G^*\mathbb{D}]$.*

Proof It is straightforward to verify that a diagram of shape $G(\mathbb{C})$ in $\mathbb{D}$, as defined in 2.3.11, is the same thing as a diagram of shape $\mathbb{C}$ in $G^*\mathbb{D}$. So this is immediate from 2.3.13. □

We note that if G has a left adjoint F, then 2.3.14 becomes a corollary of 2.3.6, since change of base along F is then left adjoint to change of base along G.

The next result is a useful extension of 1.3.18.

Lemma 2.3.15 *Let $\mathcal{S}$ be a locally cartesian closed category, and $\mathbb{C}$ an internal category in $\mathcal{S}$.*

(i) *If $\mathbb{D}$ is a locally small $\mathcal{S}$-indexed category, then so is $\mathbb{D}^{\mathbb{C}}$.*

(ii) *If $\mathbb{D}$ is an internal category in $\mathcal{S}$, then $[\mathbb{D}]^{\mathbb{C}}$ is essentially small, and may be identified with $[\mathbb{D}^{\mathbb{C}}]$, where $\mathbb{D}^{\mathbb{C}}$ is the exponential in the (cartesian closed) category* $\mathbf{Cat}(\mathcal{S})$.

Proof (i) We shall construct an enrichment of $\mathcal{D}^{\mathbb{C}}$ over $\mathcal{S}$; the extension of this to an indexed enrichment, i.e. a locally internal category structure on $\mathbb{D}^{\mathbb{C}}$, is straightforward. Let (F, ϕ) and (G, ψ) be objects of $\mathcal{D}^{\mathbb{C}}$, and let $x \colon I \to C_0$ be the object of $\mathcal{S}/C_0$ indexing morphisms $F \to G$ in $\mathcal{D}^{C_0}$, and $y \colon J \to C_1$ the object indexing morphisms $d_1^* F \to d_0^* G$ in $\mathcal{D}^{C_1}$. If we form the pullback

$$\begin{tikzcd} I' \arrow[r, "x'"] \arrow[d] & C_1 \arrow[d, "d_1"] \\ I \arrow[r, "x"] & C_0 \end{tikzcd}$$

then x' indexes morphisms $d_1^* F \to d_1^* G$ in $\mathcal{D}^{C_1}$; if we compose the generic such morphism with ψ we get a morphism $d_1^* F \to d_0^* G$, and this operation of composition induces a morphism $x' \to y$ in $\mathcal{S}/C_1$, or equivalently a morphism $x \to \Pi_{d_1} y$ in $\mathcal{S}/C_0$. Similarly, the operation of composition with ϕ yields a morphism $x \to \Pi_{d_0} y$. If we apply Π_{C_0} to each of these, we get a parallel pair of morphisms $\Pi_{C_0} x \rightrightarrows \Pi_{C_1} y$ in $\mathcal{S}$; taking the equalizer of these produces the required object of morphisms from (F, ϕ) to (G, ψ).

(ii) Given (i), to show that $[\mathbb{D}]^{\mathbb{C}}$ is essentially small we need only construct an object of $\mathcal{S}$ indexing its objects (up to isomorphism). But, by 2.3.3, any indexed functor $[\mathbb{C}] \to [\mathbb{D}]$ is canonically isomorphic to one induced by an internal functor $\mathbb{C} \to \mathbb{D}$, and we may index the latter by an appropriate subobject of $D_0^{C_0} \times D_1^{C_1}$ (specifically, an intersection of equalizers corresponding to the equations which a pair (f_0, f_1) must satisfy in order to define an internal functor). The fact that this construction, combined with that of part (i) for defining the object of morphisms of $\mathbb{D}^{\mathbb{C}}$, defines an exponential in $\mathbf{Cat}(\mathcal{S})$ is entirely straightforward. □

We remark that, for cartesian closedness of $\mathbf{Cat}(\mathcal{S})$, we need only the (proper) cartesian closedness of $\mathcal{S}$, not local cartesian closedness. On the other hand, $\mathbf{Cat}(\mathcal{S})$ is not locally cartesian closed even if $\mathcal{S}$ is: we have already observed in Section A1.5 that this fails for $\mathcal{S} = \mathbf{Set}$.

Proposition 2.3.16 *Let $\mathbb{C}$ be an internal category in $\mathcal{S}$, and let $\mathbb{D}$ be an $\mathcal{S}$-indexed category having $\mathcal{S}$-indexed products (resp. $\mathcal{S}$-indexed coproducts), as defined in Section* B1.4. *Then the category $\mathcal{D}^{\mathbb{C}}$ is comonadic (resp. monadic) over $\mathcal{D}^{C_0}$. Moreover, the indexed category $\mathbb{D}^{\mathbb{C}}$ has $\mathcal{S}$-indexed products (resp. coproducts).*

We remark in passing that we have seen the first part of this result before, in a particular case: when $\mathbb{C}$ is the internal groupoid associated to an effective equivalence relation, as in 2.3.12(c), it is just Proposition 1.5.5(i).

Proof We shall give the proof in the case when $\mathbb{D}$ has products; the other case is exactly similar. Consider the functor

$$\mathcal{D}^{C_0} \xrightarrow{d_0^*} \mathcal{D}^{C_1} \xrightarrow{\Pi_{d_1}} \mathcal{D}^{C_0};$$

we claim that this has a comonad structure. The counit is the composite

$$\Pi_{d_1} d_0^* \longrightarrow \Pi_{d_1}\Pi_{s_0} s_0^* d_0^* \cong 1$$

of which the first part is induced by the unit of $(s_0^* \dashv \Pi_{s_0})$ and the second is the coherence isomorphism arising from the fact that $d_1 s_0 = d_0 s_0 = 1_{C_0}$; the comultiplication is similarly obtained as

$$\Pi_{d_1} d_0^* \longrightarrow \Pi_{d_1}\Pi_{d_1} d_1^* d_0^* \cong \Pi_{d_1}\Pi_{d_2} d_0^* d_0^* \longrightarrow \Pi_{d_1} d_0^* \Pi_{d_1} d_0^*$$

where the final step uses the Beck–Chevalley condition for the pullback square displayed in Definition 2.3.1(a). The fact that these natural transformations satisfy the commutative diagrams for a comonad follows straightforwardly from the equations in the definition of an internal category.

Now, given an object F of $\mathcal{D}^{C_0}$, it is also straightforward to verify that a morphism $\phi\colon d_1^*F \to d_0^*F$ in $\mathcal{D}^{C_1}$ makes (F, ϕ) into a diagram of shape $\mathbb{C}$ iff its transpose $\bar{\phi}\colon F \to \Pi_{d_1} d_0^* F$ is a coalgebra structure for the comonad just described; and morphisms of diagrams similarly correspond to morphisms of coalgebras. Thus $\mathcal{D}^{\mathbb{C}}$ is isomorphic to the category of $\Pi_{d_1} d_0^*$-coalgebras.

For the second assertion, we note that for each object I of $\mathcal{S}$ we have a comonad on $\mathcal{D}^{C_0 \times I}$ whose functor part is $\Pi_{d_1 \times 1}(d_0 \times 1)^*$, and whose category of coalgebras is equivalent to $\mathcal{D}^{\mathbb{C} \times I}$. Moreover, the Beck–Chevalley condition ensures that these comonads fit together to form an indexed comonad, and that the functor part of this comonad preserves $\mathcal{S}$-indexed products. So the result follows from (the dual of) 1.4.15(ii). □

Of course, in the situation of 2.3.16, the indexed category $\mathbb{D}^{\mathbb{C}}$ also inherits $\mathcal{S}$-indexed finite limits (resp. colimits) if these are present in $\mathbb{D}$.

Corollary 2.3.17 *Let $\mathbb{C}$ be an internal category in $\mathcal{S}$, and let $\mathbb{D}$ be an $\mathcal{S}$-complete indexed category such that the fibre $\mathcal{D}^{C_0}$ is a topos (resp. a quasitopos). Then $\mathcal{D}^{\mathbb{C}}$ is a topos (resp. a quasitopos).*

Proof The functor d_0^* is cartesian because $\mathbb{D}$ is $\mathcal{S}$-complete, and Π_{d_1} is cartesian because it has a left adjoint. So this follows from A4.2.1. □

In particular, we may now deduce the analogue of A2.1.3 where **Set** is replaced by an arbitrary (quasi)topos.

Corollary 2.3.18 *If $\mathcal{S}$ is a (quasi)topos, then so is $\mathcal{S}^{\mathbb{C}}$ for any internal category $\mathbb{C}$ in $\mathcal{S}$.*

Proof Since $\mathcal{S}$ is locally cartesian closed, its canonical indexing over itself is $\mathcal{S}$-complete by 1.4.7; and the fibres of this indexed category (the slice categories $\mathcal{S}/I$) are (quasi)toposes by A2.3.2 (or its quasitopos analogue). So this is immediate from 2.3.17. □

Remark 2.3.19 In view of 2.3.12(b), Corollary 2.3.18 contains as a particular case the result, promised after A2.1.3, that $[\mathcal{C}, \mathcal{S}]$ is a topos for any finite category $\mathcal{C}$ and any topos $\mathcal{S}$. Further, if $\mathcal{S}$ is cocomplete, then $[\mathcal{C}, \mathcal{S}]$ is a topos for any small category $\mathcal{C}$.

In another direction, 2.3.16 yields the result, promised in Section B1.4, that $\mathcal{S}$-complete indexed categories have 'all small limits'.

Proposition 2.3.20 *Let $\mathbb{D}$ be an $\mathcal{S}$-complete (resp. $\mathcal{S}$-cocomplete) indexed category. Then the canonical extension of $\mathbb{D}$ to a $\mathbf{Cat}(\mathcal{S})$-indexed category, defined in 2.3.12(a), has $\mathbf{Cat}(\mathcal{S})$-indexed finite limits (resp. colimits), and its transition functors have right (resp. left) adjoints.*

Proof As previously, we do only the complete case; the cocomplete one is similar. If $f : \mathbb{C} \to \mathbb{C}'$ is an internal functor in $\mathcal{S}$, then we have a commutative diagram

$$\begin{tikzcd} \mathcal{D}^{\mathbb{C}'} \arrow[r, "f^*"] \arrow[d] & \mathcal{D}^{\mathbb{C}} \arrow[d] \\ \mathcal{D}^{C_0'} \arrow[r, "f_0^*"] & \mathcal{D}^{C_0} \end{tikzcd}$$

where the vertical arrows are forgetful functors. Since the comonads on $\mathcal{D}^{C_0}$ and $\mathcal{D}^{C_0'}$ are cartesian, by the argument in the proof of Corollary 2.3.17, these forgetful functors create finite limits; so the categories $\mathcal{D}^{\mathbb{C}}$ and $\mathcal{D}^{\mathbb{C}'}$ have finite limits, and f^* preserves them since f_0^* does. Also, since $\mathcal{D}^{\mathbb{C}'}$ has equalizers, we may lift the right adjoint Π_{f_0} of f_0^* to a right adjoint for f^*. □

We normally write $\varprojlim_f$ and $\varinjlim_f$ for the right and left adjoints, respectively, of $f^* : \mathcal{D}^{\mathbb{C}'} \to \mathcal{D}^{\mathbb{C}}$, when they exist; in the case when $\mathcal{S}$ is **Set** and $\mathbb{D}$ is the naive indexing of an ordinary complete (resp. cocomplete) category, they are familiar as the *right* and *left Kan extension* functors along f, and we shall use the same

names for them in this more general context. In particular, when $\mathbb{C}'$ is the terminal object 1 of $\mathbf{Cat}(\mathcal{S})$, so that $\mathcal{D}^{\mathbb{C}'}$ reduces to the underlying ordinary category $\mathcal{D}$ of $\mathbb{D}$, then we write $\varprojlim_{\mathbb{C}}$ and $\varinjlim_{\mathbb{C}}$ for these adjoints, and think of them as the functors assigning limits and colimits to arbitrary diagrams of type $\mathbb{C}$ in $\mathcal{D}$.

We can interpret 2.3.20 as saying that an $\mathcal{S}$-complete (resp. $\mathcal{S}$-cocomplete) indexed category is $\mathcal{A}$-complete (resp. $\mathcal{A}$-cocomplete) in the sense of 1.1.16 as an object of $\mathfrak{CAT}_{\mathcal{S}}$, where $\mathcal{A}$ is the class of all morphisms between (the canonical indexings of) internal categories in $\mathcal{S}$. However, it does *not* assert that the extension of $\mathbb{D}$ is $\mathbf{Cat}(\mathcal{S})$-complete or cocomplete: in general, the Beck–Chevalley conditions fail to hold for pullback squares in $\mathbf{Cat}(\mathcal{S})$, as we saw for the case $\mathcal{D} = \mathcal{S} = \mathbf{Set}$ in 1.4.4(b). But there is an important class of pullback squares for which they do hold:

Scholium 2.3.21 *Under the hypotheses of* 2.3.20, *the Beck–Chevalley condition holds for pullback squares of the form*

$$\begin{tikzcd}
\mathbb{B} \times \mathbb{C} \arrow[r, "f \times 1"] \arrow[d, "1 \times g"] & \mathbb{B}' \times \mathbb{C} \arrow[d, "1 \times g"] \\
\mathbb{B} \times \mathbb{C}' \arrow[r, "f \times 1"] & \mathbb{B}' \times \mathbb{C}'
\end{tikzcd}$$

in $\mathbf{Cat}(\mathcal{S})$.

Proof Consider first the case when $\mathbb{B}$ and $\mathbb{B}'$ are discrete categories, corresponding to objects B and B' of $\mathcal{S}$. The construction of the functor $\varprojlim_{1\times g} \colon \mathcal{D}^{B\times\mathbb{C}} \to \mathcal{D}^{B\times\mathbb{C}'}$ uses the functor $\Pi_{1\times g} \colon \mathcal{D}^{B\times C_0} \to \mathcal{D}^{B\times C_0'}$, the cofree and forgetful functors between these categories and the diagram categories, and the construction of an equalizer in $\mathcal{D}^{B\times\mathbb{C}'}$; but the functors $(f\times 1)^*$ commute in a suitable sense with all of these, and so the diagram

$$\begin{tikzcd}
\mathcal{D}^{B'\times\mathbb{C}} \arrow[r, "(f \times 1)^*"] \arrow[d, "\varprojlim_{1\times g}"] & \mathcal{D}^{B\times\mathbb{C}} \arrow[d, "\varprojlim_{1\times g}"] \\
\mathcal{D}^{B'\times\mathbb{C}'} \arrow[r, "(f \times 1)^*"] & \mathcal{D}^{B\times\mathbb{C}'}
\end{tikzcd}$$

commutes up to an isomorphism, which is easily seen to be the canonical natural transformation between the two composites. In other words, the functor $\varprojlim_g$ extends to an $\mathcal{S}$-indexed right adjoint for $g^* \colon \mathbb{D}^{\mathbb{C}'} \to \mathbb{D}^{\mathbb{C}}$.

In the general case, an easy extension of 2.3.13 shows that we may identify the diagram category $\mathcal{D}^{\mathbb{B}\times\mathbb{C}}$ with the category $[\mathbb{B}, [\![\mathbb{C}, \mathbb{D}]\!]]$ of indexed functors

from $\mathbb{B}$ (that is, from the indexed category $[\mathbb{B}]$) to the $\mathcal{S}$-indexed functor category $[\![\mathbb{C}, \mathbb{D}]\!]$ defined before 1.2.5. Moreover, this equivalence identifies $(f \times 1)^*$ with the functor induced by composition with f on the right, and $(1 \times g)^*$ with $[\mathbb{B}, g^*]$. Since $[\mathbb{C}, -]$ is a 2-functor $\mathfrak{CAT}_{\mathcal{S}} \to \mathfrak{CAT}$, the right adjoint of the latter is given by $[\mathbb{C}, \varprojlim_g]$; and this clearly commutes with the operation of composition with f. □

We shall also see in 2.5.11 that, at least for the case when $\mathbb{D}$ is $\mathbb{S}$ itself, we do have the Beck–Chevalley condition for comma squares in $\mathfrak{Cat}(\mathcal{S})$: thus we may say that $\mathbb{S}$ is pointwise $\mathcal{A}$-cocomplete in the sense of 1.1.16, for the class $\mathcal{A}$ mentioned earlier.

We conclude this section with the 'relative' version of A4.1.4, which is an easy consequence of 2.3.20.

Corollary 2.3.22 *Let $\mathcal{S}$ be a topos. Then, for any internal functor $f: \mathbb{C} \to \mathbb{D}$ in $\mathcal{S}$, the functor $f^*: \mathcal{S}^{\mathbb{D}} \to \mathcal{S}^{\mathbb{C}}$ is the inverse image of a geometric morphism; and in fact the assignment $\mathbb{C} \mapsto \mathcal{S}^{\mathbb{C}}$ becomes a functor* $\mathbf{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$.

Proof The canonical indexing of $\mathcal{S}$ over itself is both $\mathcal{S}$-complete and $\mathcal{S}$-cocomplete, by 1.4.7; so f^* has both left and right adjoints. The construction is clearly (pseudo-)functorial in f; and it maps the terminal object of $\mathbf{Cat}(\mathcal{S})$ to (a category isomorphic to) $\mathcal{S}$ itself. □

We note that, if we identify objects of $\mathcal{S}$ with discrete internal categories, then the functor defined in 2.3.22 extends the functor $\mathcal{S} \to \mathfrak{Top}/\mathcal{S}$ of A4.1.2. In fact we can also make it into a 2-functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$: an internal natural transformation $\alpha: f \to g$ between internal functors $\mathbb{C} \to \mathbb{C}'$ induces a natural transformation $f^* \to g^*$, whose value at an object (F, ϕ) of $\mathcal{S}^{\mathbb{C}'}$ is the morphism $\alpha^*\phi: f^*F \cong \alpha^* d_1^* F \to \alpha^* d_0^* F \cong g^* F$.

Suggestions for further reading: Betti [112, 113], Bunge [183], Pavlović [941], Street [1131].

B2.4 The Indexed Adjoint Functor Theorem

We have now assembled all the machinery necessary to prove the Adjoint Functor Theorem in the context of $\mathcal{S}$-indexed categories, where $\mathcal{S}$ is an arbitrary cartesian category. It is well known that there are two versions ('General' and 'Special') of this theorem, and both of them have indexed analogues (see [933]); but we shall prove only the 'Special' version here – it is this version that tends to be more useful in applications.

We begin by introducing the indexed version of the notion of separating family (cf. A1.2.4).

Definition 2.4.1 Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category. An object G of $\mathcal{C}^I$ (for some object I of $\mathcal{S}$) is called a *separating family* for $\mathbb{C}$ if, given any $J \in \text{ob}\ \mathcal{S}$ and any

$f, g \colon A \rightrightarrows B$ in $\mathcal{C}^J$ with $f \neq g$, there exists a span

$$I \xleftarrow{\ x\ } K \xrightarrow{\ y\ } J$$

in $\mathcal{S}$ and a morphism $h \colon x^*(G) \to y^*(A)$ in $\mathcal{C}^K$ with $y^*(f)h \neq y^*(g)h$. In the particular case when I is the terminal object 1 of $\mathcal{S}$, we call G a *separator.*

Examples 2.4.2 (*a*) For the naive indexing of an ordinary category $\mathcal{C}$ over **Set**, a separating family is exactly what we meant by that term in Section A1.2. For if $(G_i)_{i \in I}$ is a separating family in the usual sense, and we are given two J-indexed families $(f_j, g_j \colon A_j \rightrightarrows B_j)_{j \in J}$ which are not equal, then there exists $j \in J$ with $f_j \neq g_j$, and there exists $i \in I$ and $h \colon G_i \to A_j$ with $f_j h \neq g_j h$, so we have a span of the required form with $K = 1$. The converse is similar.

(*b*) For the canonical indexing $\mathbb{S}$ of $\mathcal{S}$ over itself, the terminal object 1 of $\mathcal{S} \cong \mathcal{S}^1$ is always a separator. For if $u, v \colon y \rightrightarrows z$ are two unequal maps in $\mathcal{S}/J$, then we take $K = \operatorname{dom} y$ and the span

$$1 \xleftarrow{\ K\ } K \xrightarrow{\ y\ } J\ ,$$

together with the diagonal map $h \colon K \to K \times_J K$ regarded as a morphism $K^*(1) \to y^*(y)$ in $\mathcal{S}/K$. Then $y^*(u)h$ and $y^*(v)h$ are simply the graphs of u and v, so they are not equal. However, this example does not always have a coseparating family (though it does if $\mathcal{S}$ is a topos; cf. 3.1.13 below).

(*c*) A small (that is, internal) category $\mathbb{C}$ always has a separating family, namely the family of all objects of $\mathbb{C}$ (that is, the identity morphism 1_{C_0} in ob $\mathcal{C}^{C_0}$). Once again, this is easy to verify: we take $K = J$, $y = 1_J$ and x to be the morphism $J \to C_0$ which is the domain of the given parallel pair in $\mathcal{C}^J$, together with the identity morphism $x^*(1_{C_0}) \to x$.

(*d*) Let $L \colon \mathbb{C} \to \mathbb{D}$ be an indexed functor having an indexed right adjoint F which is faithful (i.e. such that F^I is faithful for each I). Then if $G \in$ ob $\mathcal{C}^I$ is a separating family for $\mathbb{C}$, it is easy to verify that $L^I(G)$ is a separating family for $\mathbb{D}$. For if $f, g \colon A \rightrightarrows B$ are unequal maps in $\mathcal{D}^J$, we have $F^J(f) \neq F^J(g)$; so we can find a span (x, y) in $\mathcal{S}$ and a morphism $h \colon x^*(G) \to y^* F^J(A) \cong F^K y^*(A)$ having unequal composites with $F^K y^*(f)$ and $F^K y^*(g)$, and then the transpose $\overline{h} \colon x^* L^I(G) \cong L^K x^*(G) \to y^*(A)$ is the morphism we seek. In particular, if $\mathbb{C}$ has a separating family, then any indexed reflective subcategory of $\mathbb{C}$ has a separating family.

In constructing further examples of separating families, we shall find the following lemma useful.

Lemma 2.4.3 *Suppose $\mathbb{C}$ is $\mathcal{S}$-cocomplete and locally small. Then $G \in$ ob $\mathcal{C}^I$ is a separating family for $\mathbb{C}$ iff, for every $J \in$ ob $\mathcal{S}$ and every $A \in$ ob $\mathcal{C}^J$, there is a span*

$$I \xleftarrow{\ x\ } K \xrightarrow{\ y\ } J$$

in $\mathcal{S}$ and an epimorphism $e \colon \Sigma_y x^*(G) \twoheadrightarrow A$ in $\mathcal{C}^J$.

Proof Suppose the condition holds. Then, for any $f, g\colon A \rightrightarrows B$ in $\mathcal{C}^J$ with $f \neq g$, we have $fe \neq ge$, so we obtain $y^*(f)h \neq y^*(g)h$ where $h\colon x^*(G) \to y^*(A)$ is the transpose of e across the adjunction $(\Sigma_y \dashv y^*)$.

Conversely, suppose G is a separating family. Given A, let $(x, y)\colon K \to I \times J$ be the object indexing morphisms $G \to A$, and $h\colon x^*(G) \to y^*(A)$ the generic such morphism in $\mathcal{C}^K$. Now, given $f, g\colon A \rightrightarrows B$ in $\mathcal{C}^J$ with $f \neq g$, the definition of separating family says that we can find $(x', y')\colon K' \to I \times J$ and $h'\colon x'^*(G) \to y'^*(A)$ with $y'^*(f)h' \neq y'^*(g)h'$. But h' corresponds to a morphism $z\colon K' \to K$ in $\mathcal{S}$ such that (modulo canonical isomorphisms) $z^*(h) = h'$; so this means that $y^*(f)h \neq y^*(g)h$ for any such pair (f, g). Hence, if we transpose h across the adjunction $(\Sigma_y \dashv y^*)$, we obtain $e\colon \Sigma_y x^*(G) \to A$ such that $fe \neq ge$ whenever $f \neq g$, i.e. such that e is an epimorphism. □

The possession of a separating family is not in general preserved under change of base, even when the object indexing the separating family lies in the image of the change-of-base functor. (Counterexamples can be given using 2.4.2(*b*): we shall see in Section B3.1 that there are geometric morphisms $p\colon \mathcal{E} \to \mathcal{S}$ for which the indexing of $\mathcal{E}$ over $\mathcal{S}$ induced by p^* does not have a separating family.) The following result, though rather special, will be needed in the proof of the Adjoint Functor Theorem.

Lemma 2.4.4 *Let $\mathcal{S}$ be a category with finite products, I an object of $\mathcal{S}$. Then the change-of-base functor $\Sigma_I^*\colon \mathfrak{Cat}_{\mathcal{S}} \to \mathfrak{Cat}_{\mathcal{S}/I}$ preserves the property of having a separating family.*

Proof Let $\mathbb{C}$ be an $\mathcal{S}$-indexed category having a separating family $G \in \text{ob}\, \mathcal{C}^J$. We shall show that the object $\pi_1^*(G)$ of $\mathcal{C}^{J \times I} = (\Sigma_I^* \mathcal{C})^{I^*(J)}$ is a separating family for $\Sigma_I^* \mathbb{C}$. Given an object $u\colon K \to I$ of $\mathcal{S}/I$ and an unequal parallel pair $f, g\colon A \rightrightarrows B$ in $\mathcal{C}^K$, we have a span

$$J \xleftarrow{\;x\;} L \xrightarrow{\;y\;} K$$

in $\mathcal{S}$ and a morphism $h\colon x^*(G) \to y^*(A)$ such that $y^*(f)h \neq y^*(g)h$. Then $z = (x, uy)\colon L \to J \times I$ is a morphism $uy \to I^*(J)$ in $\mathcal{S}/I$, and we have $z^*\pi_1^*(G) \cong x^*(G)$; so

$$I^*(J) \xleftarrow{\;z\;} uy \xrightarrow{\;y\;} u$$

is the required span in $\mathcal{S}/I$. □

It is customary, in textbooks on category theory, to present the Adjoint Functor Theorem as a criterion for the existence of left adjoints. However, we shall probably make more use of it (insofar as we use it at all) in constructing right adjoints; and so we shall present the dual of the 'usual' form of the theorem. Like other authors, we shall begin by considering a special case – the criterion for

existence of a terminal object – and then deduce the general case by consideration of comma categories.

Proposition 2.4.5 (The Indexed Special Terminal-Object Theorem) *Let $\mathcal{S}$ be a cartesian category, and let $\mathbb{C}$ be a cocomplete, locally small, well-copowered $\mathcal{S}$-indexed category with a separating family. Then $\mathbb{C}$ has an $\mathcal{S}$-indexed terminal object (that is, the categories $\mathcal{C}^I$ have terminal objects, and the transition functors x^* preserve them).*

Proof It suffices to prove that $\mathcal{C}^1$ has a terminal object; since the transition functors have left adjoints, they will preserve the terminal object if it exists.

Let $G \in \mathrm{ob}\, \mathcal{C}^I$ be a separating family for $\mathcal{C}$, and write A for the object $\Sigma_I(G)$ of $\mathcal{C}^1$. Let Q be the object indexing the quotients of A in $\mathbb{C}$, and $e\colon Q^*(A) \twoheadrightarrow A'$ the generic quotient of A in $\mathcal{C}^Q$. In $\mathcal{C}^{Q\times Q}$ we have a generic pair of quotients $\pi_1^*(e)$ and $\pi_2^*(e)$; we may form their cointersection (= pushout), and classify this by a third morphism $z\colon Q \times Q \to Q$ in $\mathcal{S}$. Let $Q_1 \rightarrowtail Q$ be the equalizer of z and π_2; then (by the dual of an argument familiar for subobjects; cf. A1.6.3) $\mathbb{Q} = (Q_1 \rightrightarrows Q)$ has the structure of an internal poset in $\mathcal{S}$. Moreover, the object A' of $\mathcal{C}^Q$ acquires the structure of an internal diagram (A', α) in $\mathcal{C}^{\mathbb{Q}}$, in an obvious way. And the epimorphism $Q^*(A) \twoheadrightarrow A'$ becomes a morphism from the constant diagram $\mathbb{Q}^*(A)$ to (A', α) – which is again epic, since the forgetful functor $\mathcal{C}^{\mathbb{Q}} \to \mathcal{C}^Q$ is faithful.

Let T be the object $\varinjlim_{\mathbb{Q}}(A', \alpha)$ of $\mathcal{C}^1$. Since $\mathbb{Q}$ is connected (because it has an initial object, corresponding to the largest quotient $1_A\colon A \twoheadrightarrow A$), we have $\varinjlim_{\mathbb{Q}} \mathbb{Q}^*(A) \cong A$; and since $\varinjlim_{\mathbb{Q}}$, being a left adjoint, preserves epimorphisms, we have an epimorphism $h\colon A \twoheadrightarrow T$ in $\mathcal{C}^1$. Moreover, the transpose of the canonical epimorphism $\Sigma_Q(A') \twoheadrightarrow T$ (the 'colimiting cone') is a morphism $A' \to Q^*(T)$ in $\mathcal{C}^Q$, which is easily seen to make the diagram

$$\begin{tikzcd}
Q^*(A) \arrow[r, two heads] \arrow[dr, "Q^*(h)"'] & A' \arrow[d] \\
 & Q^*(T)
\end{tikzcd}$$

commute; so $h\colon A \twoheadrightarrow T$ is the unique smallest quotient of A.

We claim that T is a terminal object of $\mathcal{C}^1$. Suppose B is any object of $\mathcal{C}^1$; then by 2.4.3 we have a morphism $x\colon J \to I$ in $\mathcal{S}$ and an epimorphism $\Sigma_J x^*(G) \twoheadrightarrow B$ in $\mathcal{C}^1$. But $\Sigma_J x^*(G) \cong \Sigma_I \Sigma_x x^*(G)$, so we may form the pushout

$$\begin{tikzcd}
\Sigma_J x^*(G) \arrow[r, two heads] \arrow[d] & B \arrow[d] \\
\Sigma_I(G) \arrow[r, two heads] & C
\end{tikzcd}$$

where the left vertical map is induced by the counit of $(\Sigma_x \dashv x^*)$. Since C is a quotient of $\Sigma_I(G) = A$, it admits a morphism to T; so B admits a morphism to T. But if we had two parallel morphisms $f, g\colon B \rightrightarrows T$, then their coequalizer $T \twoheadrightarrow T'$ would be a quotient of T, and hence a quotient of A contained in T. So $T \twoheadrightarrow T'$ must be an isomorphism, and hence $f = g$. $\square$

Theorem 2.4.6 (The Indexed Special Adjoint Functor Theorem) *Let $\mathcal{S}$ be a cartesian category; let $\mathbb{C}$ and $\mathbb{D}$ be $\mathcal{S}$-indexed categories which are locally small and cocomplete, and suppose further that $\mathbb{C}$ is well-copowered and has a separating family. Then an indexed functor $F\colon \mathbb{C} \to \mathbb{D}$ has an indexed right adjoint iff it is cocontinuous.*

Proof One direction is just (the dual of) 1.4.14. Conversely, suppose F is cocontinuous. To construct an indexed right adjoint R for it, it suffices in fact to construct an ordinary right adjoint $R^I\colon \mathcal{D}^I \to \mathcal{C}^I$ for each F^I; for the fact that the F^I commute up to isomorphism with the left adjoints Σ_x of the transition functors will ensure that the R^I form an indexed functor (cf. the proof of 1.4.14).

Consider first an object B of $\mathcal{C}^1$. To construct $R^1(B)$, we need to find a terminal object of the (ordinary) comma category $(F^1 \downarrow B)$, as is well known. But this comma category is the fibre over 1 of an 'indexed comma category' which we shall denote by $(\mathbb{F} \downarrow B)$, and whose fibre over I is the ordinary comma category $(F^I \downarrow I^*(B))$; the transition functors are induced in the obvious way by those of $\mathbb{C}$. We shall show that $(\mathbb{F} \downarrow B)$ satisfies the hypotheses of 2.4.5, and so has an indexed terminal object; in particular, $(F^1 \downarrow B)$ has a terminal object.

First we must verify that $(\mathbb{F} \downarrow B)$ is cocomplete. The existence of finite colimits in the fibres $(F^I \downarrow I^*(B))$, and the fact that the transition functors preserve them, follows straightforwardly from the fact that the $\mathcal{C}^I$ have finite limits, and the F^I and the transition functors of $\mathbb{C}$ both preserve them. So we need only construct the left adjoints Σ_x for the transition functors of $(\mathbb{F} \downarrow B)$, and verify that they satisfy the Beck–Chevalley condition. Given $x\colon I \to J$ in $\mathcal{S}$ and an object $(A, g\colon F^I(A) \to I^*(B))$ of $(F^I \downarrow I^*(B))$, we define $\Sigma_x(A, g)$ to be $(\Sigma_x(A), h)$, where h is the composite

$$F^J\Sigma_x(A) \cong \Sigma_x F^I(A) \xrightarrow{\Sigma_x(g)} \Sigma_x I^*(B) \cong \Sigma_x x^* J^*(B) \longrightarrow J^*(B)$$

(the last arrow being the counit of $(\Sigma_x \dashv x^*)$). It is straightforward to verify that this object has the right universal property, and that the Beck–Chevalley condition holds.

Next, we verify local smallness. Given two objects (A, g) and (A', g') of $(F^I \downarrow I^*(B))$ and $(F^J \downarrow J^*(B))$ respectively, we have an object $(x, y)\colon K \to I \times J$ of $\mathcal{S}/I \times J$ indexing morphisms from A to A' in $\mathbb{C}$. Let $f\colon x^*(A) \to y^*(A')$ be

the generic such morphism; then in $\mathcal{D}^K$ we have a (non-commuting) triangle

$$\begin{tikzcd} F^K x^*(A) \arrow[r, "x^*(g)"] \arrow[d, "F^K(f)"'] & K^*(B) \\ F^K y^*(A') \arrow[ur, "y^*(g')"'] & \end{tikzcd}$$

(where we have suppressed the names of canonical isomorphisms). But equality of morphisms is definable in $\mathbb{D}$, by 1.3.15; so we have a monomorphism $z\colon K' \rightarrowtail K$ which is universal among morphisms $u\colon L \to K$ such that u^* maps the above diagram to a commuting triangle. Then it is easy to see that $(xz, yz)\colon K' \to I \times J$ indexes morphisms from (A, g) to (A', g') in $(\mathbb{F} \downarrow B)$.

The argument for well-copoweredness is similar. For any I, the forgetful functor $(F^I \downarrow I^*(B)) \to \mathcal{C}^I$ preserves and reflects epimorphisms, since F^I preserves them; so, given an object (A, g) of $(F^I \downarrow I^*(B))$, we have simply to find an object indexing those quotients $A \twoheadrightarrow A'$ of A for which g factors through $FA \twoheadrightarrow FA'$. If $x\colon J \to I$ indexes quotients of A in $\mathbb{C}$, this will (in general) not be the case for the generic quotient $e\colon x^*(A) \twoheadrightarrow A'$ in $\mathcal{C}^J$; but we can form the pushout

$$\begin{tikzcd} F^J x^*(A) \arrow[r, "x^*(g)"] \arrow[d, two heads, "F^J(e)"] & J^*(B) \arrow[d, "h"] \\ F^J(A') \arrow[r] & C \end{tikzcd}$$

and then, using the fact that invertibility is definable in $\mathbb{D}$ (1.3.15 again), form the subobject $y\colon J' \rightarrowtail J$ which is universal among morphisms z such that $z^*(h)$ is an isomorphism. Since the transition functors of $\mathbb{D}$ preserve pushouts, it follows easily that $xy\colon J' \to I$ indexes quotients of (A, g) in $(\mathbb{F} \downarrow B)$.

Finally, we must verify that $(\mathbb{F} \downarrow B)$ has a separating family. Let $G \in \mathrm{ob}\, \mathcal{C}^I$ be a separating family for $\mathbb{C}$, and let $(x, J)\colon J \to I \times 1$ be the object indexing morphisms from $F^I(G)$ to B in $\mathbb{D}$. Let $g\colon F^J x^*(G) \cong x^* F^I(G) \to J^*(B)$ be the generic such morphism; then we claim that $(x^*(G), g)$ is a separating family for $(\mathbb{F} \downarrow B)$. To see this, let (A, h) be an object of $(F^K \downarrow K^*(B))$ for some K; then by 2.4.3 we have a span

$$I \xleftarrow{\;y\;} L \xrightarrow{\;z\;} K$$

in $\mathcal{S}$ and a morphism $f\colon y^*(G) \to z^*(A)$ whose transpose $\Sigma_z y^*(G) \to A$ is an epimorphism. Now the composite

$$y^* F^I(G) \cong F^L y^*(G) \xrightarrow{F^L(f)} F^L z^*(A) \cong z^* F^K(A) \xrightarrow{z^*(h)} z^* K^*(B) \cong L^*(B)$$

defines a morphism $w\colon L \to J$ such that $xw = y$ and the composite above is (isomorphic to) $w^*(g)$; so we have $w^*(x^*(G), g) \cong (y^*(G), w^*(g))$ and f becomes a morphism from this object to $z^*(A, h)$, whose transpose is epimorphic.

Thus we may apply 2.4.5 to $(\mathbb{F} \downarrow B)$, to obtain a terminal object $(R^1(B), \epsilon_B)$ of $(F^1 \downarrow B)$. If we do this for all objects B of $\mathcal{D}^1$, it is well known that R^1 becomes the object-map of a functor $\mathcal{D}^1 \to \mathcal{C}^1$, such that ϵ is the counit of an adjunction $(F^1 \dashv R^1)$. To obtain the right adjoint R^I of F^I, for a general object I of $\mathcal{S}$, we simply repeat the whole argument in the context of $\mathcal{S}/I$-indexed categories, using the fact that the change-of-base functor along $\Sigma_I\colon \mathcal{S}/I \to \mathcal{S}$ preserves all the properties used in the hypotheses of the theorem, by 1.4.9, 2.2.6(ii), 2.2.7 and 2.4.4. □

Remark 2.4.7 We have seen in 2.4.2(c) that a small category $\mathbb{C}$ always has a separating family. It is also locally small, as we observed before 2.3.5; it need not be well-copowered in the sense in which we defined this notion in 1.3.14, but it is easy to see that if $\mathbb{C}$ has $\mathcal{S}$-indexed pushouts then the concept of an individual epimorphism in $\mathbb{C}$ with given domain (as opposed to an isomorphism class of such epimorphisms, i.e. a quotient object) is comprehensible, since a morphism is epic iff the two morphisms in its cokernel-pair are equal, and $\mathbb{C}$ has definable equality by 1.3.15. An examination of the proofs of 2.4.5 and 2.4.6 will show that this is sufficient to make them work; so we deduce that *if $\mathbb{C}$ is an $\mathcal{S}$-cocomplete small (i.e. internal) category in $\mathcal{S}$, then every cocontinuous functor from $\mathbb{C}$ to a locally small cocomplete $\mathcal{S}$-indexed category $\mathbb{D}$ has a right adjoint.*

As an application of 2.4.7, we prove a result promised after 2.3.9.

Corollary 2.4.8 *Suppose $\mathcal{S}$ is (properly) cartesian closed. Then an internal category $\mathbb{C}$ in $\mathcal{S}$ is $\mathcal{S}$-complete iff it is $\mathcal{S}$-cocomplete.*

Proof Suppose $\mathbb{C}$ is $\mathcal{S}$-cocomplete. To prove that $\mathbb{C}$ is $\mathcal{S}$-complete, it suffices to construct $\mathcal{S}$-indexed right adjoints for the diagonal functors $\Delta\colon \mathbb{C} \to [\mathcal{D}, \mathbb{C}]$ for each finite category $\mathcal{D}$, and for the indexed transition functors $x^*\colon \mathbb{C}^I \to \mathbb{C}^J$ induced by each morphism $x\colon J \to I$ of $\mathcal{S}$. (The Beck–Chevalley condition for the right adjoints to the x^* follows from that for their left adjoints, by 1.4.6.) It is easy to see that $[\mathcal{D}, \mathbb{C}]$ is cocomplete if $\mathbb{C}$ is (the left adjoints to its transition functors being defined 'pointwise') and that Δ is cocontinuous; also, $[\mathcal{D}, \mathbb{C}]$ is small (and hence locally small) if $\mathbb{C}$ is, since we can construct its object of objects and object of morphisms using appropriate finite limits in $\mathcal{S}$. Similarly, $\mathbb{C}^I$ is $\mathcal{S}$-cocomplete if $\mathbb{C}$ is, by 1.4.9; and the proof of this fact again makes it obvious that the indexed transition functors x^* are cocontinuous. The point at which we use the cartesian closedness of $\mathcal{S}$ is to ensure that $\mathbb{C}^I$ is small if $\mathbb{C}$ is: of course, its object of objects and object of morphisms are the exponentials $(C_0)^I$ and $(C_1)^I$. So the existence of the required right adjoints follows from 2.4.7. □

Suggestion for further reading: Paré & Schumacher [933].

B2.5 Discrete opfibrations

The remaining three sections of this chapter are devoted to further study of diagram categories, as defined in 2.3.11, which will be needed for the proof of Diaconescu's Theorem in Section B3.2. From now on, we shall tend to use the notation $[\mathbb{C}, \mathcal{D}]$ rather than $\mathcal{D}^{\mathbb{C}}$ for diagram categories, since we are thinking of them mainly as 'hom-categories' in the 2-category $\mathfrak{Cat}_{\mathcal{S}}$, rather than as the '$\mathbb{C}$-component' of an indexing of $\mathcal{D}$ over $\mathbf{Cat}(\mathcal{S})$.

In this section we investigate an alternative way of looking at the diagram categories $[\mathbb{C}, \mathcal{S}]$, where $\mathbb{C}$ is an internal category in a cartesian category $\mathcal{S}$. We begin with a useful notational convention regarding pullbacks over C_0. Since C_1 comes equipped with two morphisms to C_0, the usual 'fibre product' notation $C_1 \times_{C_0} A$ (where $A \to C_0$ is another object of $\mathcal{S}/C_0$) is potentially ambiguous; but, if we regard d_0 and d_1 as proceeding respectively 'leftwards' and 'rightwards' from C_1, as in the picture

$$\begin{tikzcd} & C_1 \arrow[dl, "d_0"'] \arrow[dr, "d_1"] & \\ C_0 & & C_0 \end{tikzcd} ,$$

then it makes sense to agree that (unless explicitly stated otherwise) when C_1 appears on the left of the symbol $\times_{C_0}$ it denotes the object d_1 of $\mathcal{S}/C_0$, and when it appears on the right it denotes the object d_0. Thus, for example, $C_1 \times_{C_0} C_1$ is another name for C_2, since we have a pullback square

$$\begin{tikzcd} C_2 \arrow[r, "d_2"] \arrow[d, "d_0"] & C_1 \arrow[d, "d_0"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd} ,$$

and more generally the n-fold fibre product $C_1 \times_{C_0} C_1 \times_{C_0} \cdots \times_{C_0} C_1$ denotes C_n. (Accordingly, extending our convention, $(-) \times_{C_0} C_n$ signifies that the structure map $C_n \to C_0$ is to be taken as $d_0^1 d_0^2 \cdots d_0^n$, and $C_n \times_{C_0} (-)$ that it is to be taken as $d_1^1 d_2^2 \cdots d_n^n$. Of course, there are $(n-2)$ other morphisms from C_n to C_0 built into the structure of $\mathbb{C}$, but we shall rarely need to use these when forming pullbacks.)

By definition, a diagram of shape $\mathbb{C}$ in $\mathbb{S}$ consists of an object $(f : F \to C_0)$ of $\mathcal{S}/C_0$, equipped with a morphism

$$\begin{tikzcd} C_1 \times_{C_0} F \arrow[rr, "\phi"] \arrow[dr, "\pi_1"'] & & F \times_{C_0} C_1 \arrow[dl, "\pi_2"] \\ & C_1 & \end{tikzcd}$$

in $\mathcal{S}/C_1$ which is unital and associative. Since $\mathbb{S}$ has $\mathcal{S}$-indexed coproducts, we can equivalently regard the structure as a morphism $\overline{\phi} = \pi_1\phi : C_1 \times_{C_0} F \to F$ making

$$\begin{tikzcd} C_1 \times_{C_0} F \arrow[r, "\overline{\phi}"] \arrow[d, "\pi_1"] & F \arrow[d, "f"] \\ C_1 \arrow[r, "d_0"] & C_0 \end{tikzcd}$$

commute (we think of this as a *left action* of $\mathbb{C}$ on F); and, of course, the projection $\pi_2 : C_1 \times_{C_0} F \to F$ makes

$$\begin{tikzcd} C_1 \times_{C_0} F \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & F \arrow[d, "f"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd}$$

commute. Further, the unital condition on ϕ says that

$$F \xrightarrow{(s_0 f, 1_F)} C_1 \times_{C_0} F \xrightarrow{\overline{\phi}} F$$

is the identity, and the associativity says that the two composites

$$\begin{tikzcd} C_1 \times_{C_0} C_1 \times_{C_0} F \arrow[r, shift left, "1 \times \overline{\phi}"] \arrow[r, shift right, "d_1 \times 1"'] & C_1 \times_{C_0} F \arrow[r, "\overline{\phi}"] & F \end{tikzcd}$$

are equal. It follows that if we define $F_0 = F$ and $F_n = C_n \times_{C_0} F$ for $n \geq 1$, then $n \mapsto F_n$ has the structure of an internal category $\mathbb{F}$, with $d_i^n : F_n \to F_{n-1}$ taken

to be $d_i^n \times 1_F : C_n \times_{C_0} F \to C_{n-1} \times_{C_0} F$ for $0 \leq i < n$, d_n^n taken to be

$$C_n \times_{C_0} F \cong C_{n-1} \times_{C_0} C_1 \times_{C_0} F \xrightarrow{1 \times \overline{\phi}} C_{n-1} \times_{C_0} F \ ,$$

and s_j^n similarly taken to be $s_j^n \times 1_F$. Moreover, if we define $f_0 = f$ and $f_n = \pi_1 : C_n \times_{C_0} F \to C_n$ for $n \geq 1$, we obtain an internal functor $f : \mathbb{F} \to \mathbb{C}$.

Of course, what is going on here is just the internal version of the Grothendieck construction which we met in A1.1.7, and again (in a more general setting) in Section B1.3. Accordingly, we now introduce the fundamental definition of this section:

Definition 2.5.1 A *discrete opfibration* in a cartesian category $\mathcal{S}$ is an internal functor $f : \mathbb{F} \to \mathbb{C}$ for which

$$\begin{tikzcd} F_1 \arrow[r, "d_1"] \arrow[d, "f_1"] & F_0 \arrow[d, "f_0"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd}$$

is a pullback. Similarly, we call f a *discrete fibration* if

$$\begin{tikzcd} F_1 \arrow[r, "d_0"] \arrow[d, "f_1"] & F_0 \arrow[d, "f_0"] \\ C_1 \arrow[r, "d_0"] & C_0 \end{tikzcd}$$

is a pullback.

Internalizing 1.3.11(iii), we have

Lemma 2.5.2 *Let $f : \mathbb{C} \to \mathbb{D}$ and $g : \mathbb{D} \to \mathbb{E}$ be internal functors in $\mathcal{S}$, such that g is a discrete opfibration. Then f is a discrete opfibration iff the composite gf is.*

Proof This follows from well-known properties of pullbacks. □

Proposition 2.5.3 *Let* **doFib**$(\mathcal{S})$ *denote the category of internal categories in $\mathcal{S}$ and discrete opfibrations between them. Then, for any $\mathbb{C}$, the slice category* **doFib**$(\mathcal{S})/\mathbb{C}$ *is equivalent to* $[\mathbb{C}, \mathcal{S}]$.

Proof The discussion before Definition 2.5.1 shows how each diagram (F, ϕ) of type $\mathbb{C}$ in $\mathcal{S}$ gives rise to a discrete opfibration $\mathbb{F} \to \mathbb{C}$. Conversely, a discrete opfibration $f : \mathbb{F} \to \mathbb{C}$ may be reconstructed, up to canonical

isomorphism, from the object $f_0 \colon F_0 \to C_0$ of $\mathcal{S}/C_0$, together with the morphism $d_0 \colon C_1 \times_{C_0} F_0 \cong F_1 \to F_0$, and the latter is readily seen to be a (unital and associative) left action of $\mathbb{C}$ on F_0. It is further easy to see that morphisms of diagrams $(F, \phi) \to (G, \psi)$ correspond under these constructions to internal functors $\mathbb{F} \to \mathbb{G}$ over $\mathbb{C}$; but by 2.5.2 all such functors are discrete opfibrations. □

Of course, discrete fibrations over $\mathbb{C}$ similarly correspond to diagrams of shape $\mathbb{C}^{\mathrm{op}}$ in $\mathcal{S}$ (which we may further identify with objects $(f \colon F \to C_0)$ of $\mathcal{S}/C_0$ equipped with a *right action* $F \times_{C_0} C_1 \to F$). It is also possible to define the notion of (internal) *split (op)fibration* in an arbitrary cartesian category $\mathcal{S}$ – as we remarked in Section B1.3, the theory of split fibrations is essentially algebraic – and to show that split opfibrations over $\mathbb{C}$ correspond both to diagrams of shape $\mathbb{C}$ in $\mathbf{Cat}(\mathcal{S})$, and to internal categories in $[\mathbb{C}, \mathcal{S}]$. However, we shall not pursue the details here.

Instead, we next consider some examples of discrete opfibrations which will be of importance in our subsequent discussions.

Examples 2.5.4 (*a*) If $\mathbf{1}$ denotes the terminal object of $\mathbf{Cat}(\mathcal{S})$, it is easy to see that the unique internal functor $\mathbb{F} \to 1$ is a discrete opfibration iff $\mathbb{F}$ is a discrete category. Thus the equivalence $\mathbf{doFib}(\mathcal{S})/1 \simeq \mathcal{S}^1$ of 2.5.3 identifies objects of $\mathcal{S}$ with discrete internal categories, as we did in 2.3.12(*a*). More generally, for any internal category $\mathbb{C}$ and any object I of $\mathcal{S}$ (regarded as a discrete internal category), the projection $\mathbb{C} \times I \to \mathbb{C}$ is both a discrete fibration and a discrete opfibration; it is easy to see that this corresponds to the *constant diagram* of shape $\mathbb{C}$ with value I, i.e. the object $C_0^* I$ of $\mathcal{S}/C_0$ equipped with the structure map

$$d_1^* C_0^* I \cong C_1^* I \xrightarrow{1} C_1^* I \cong d_0^* C_0^* I \ .$$

(*b*) For any internal category $\mathbb{C}$ in $\mathcal{S}$ (or, more generally, for any simplicial object), the *shift* or *décalage* $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$ of $\mathbb{C}$ is the simplicial object obtained from $\mathbb{C}$ by shifting dimensions by 1 and dropping the zeroth face and degeneracy maps: that is, $\mathrm{Dec}_1(\mathbb{C})_n$ is C_{n+1}, $d_i^n \colon \mathrm{Dec}_1(\mathbb{C})_n \to \mathrm{Dec}_1(\mathbb{C})_{n-1}$ is defined to be $d_{i+1}^{n+1} \colon C_{n+1} \to C_n$, and similarly for the s_j^n. It is easy to check that the simplicial identities for $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$ follow from those for $\mathbb{C}$, and that $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$ is an internal category if $\mathbb{C}$ is. Moreover, the omitted face maps $d_0^{n+1} \colon C_{n+1} \to C_n$ form the components of a morphism of simplicial objects $\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \to \mathbb{C}$, which is a discrete opfibration if $\mathbb{C}$ is an internal category, since

$$\begin{tikzcd} C_2 \arrow[r, "d_2"] \arrow[d, "d_0"] & C_1 \arrow[d, "d_0"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd}$$

is a pullback. The corresponding diagram of shape $\mathbb{C}$ in $\mathcal{S}$ is of course the object $d_0\colon C_1 \to C_0$ of $\mathcal{S}/C_0$, equipped with the left action $d_1 : C_1 \times_{C_0} C_1 \cong C_2 \to C_1$. We shall also write $\mathbb{D}\mathrm{ec}^1(\mathbb{C})$ for the simplicial object with the same objects of n-simplices as $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$, but with the *last* face and degeneracy maps thrown away; i.e. $d_i^n : \mathrm{Dec}^1(\mathbb{C})_n \to \mathrm{Dec}^1(\mathbb{C})_{n-1}$ is defined to be $d_i^{n+1} : C_{n+1} \to C_n$, and so on. In this case, when $\mathbb{C}$ is an internal category, the omitted face maps $d_{n+1}^{n+1} : C_{n+1} \to C_n$ form a discrete fibration $\mathbb{D}\mathrm{ec}^1(\mathbb{C}) \to \mathbb{C}$. And we shall write $\mathbb{D}\mathrm{ec}^n$ (resp. $\mathbb{D}\mathrm{ec}_n$) for the nth iterate of the functor $\mathbb{D}\mathrm{ec}^1$ (resp. $\mathbb{D}\mathrm{ec}_1$) from $[\mathbf{\Delta}^{\mathrm{op}}, \mathcal{S}]$ to itself, i.e. the operation which shifts dimensions by n and throws away the last (resp. first) n face and degeneracy maps.

(c) More generally than (*b*), if we are given any object $x\colon I \to C_0$ of $\mathcal{S}/C_0$, we may construct an internal category $\mathbb{R}(x)$ equipped with a discrete opfibration $r\colon \mathbb{R}(x) \to \mathbb{C}$, by setting $R(x)_n = C_{n+1} \times_{C_0} I$, with $r_n = d_0^{n+1}\pi_1 : C_{n+1} \times_{C_0} I \to C_n$, and the face and degeneracy maps of $\mathbb{R}(x)$ induced by those of $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$. (Thus $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$ itself is $\mathbb{R}(1_{C_0})$.) The letter $\mathbb{R}$ stands for 'representable': if $\mathcal{S} = \mathbf{Set}$ and $x\colon 1 \to C_0$ is a particular object of a small category $\mathbb{C}$, then $\mathbb{R}(x)$ is (the discrete opfibration corresponding to) the representable functor $\mathbb{C}(x, -)\colon \mathbb{C} \to \mathbf{Set}$. In general, recalling from 2.3.16 that $\mathcal{S}^{\mathbb{C}}$ is monadic over $\mathcal{S}/C_0$ (the functor part of the monad being $\Sigma_{d_0} d_1^* = C_1 \times_{C_0} (-)$), it is easy to see that $\mathbb{R}$ is the free functor $\mathcal{S}/C_0 \to [\mathbb{C}, \mathcal{S}]$, i.e. the left adjoint to the forgetful functor. Note also that all the structure maps of $\mathbb{R}(x)$ are morphisms over I; thus we may regard it equivalently as a diagram of shape $\mathbb{C}$ in $\mathcal{S}/I$, or as an I-indexed family of objects of the $\mathcal{S}$-indexed category $\mathbb{S}^{\mathbb{C}}$ (cf. 2.3.12(*a*)).

Now suppose that $\mathcal{S}$ has coequalizers of reflexive pairs (cf. A1.2.10). For any internal category $\mathbb{C}$ in $\mathcal{S}$ we define $\pi_0\mathbb{C}$, the *object of components* of $\mathbb{C}$, to be the coequalizer of $(d_0, d_1)\colon C_1 \rightrightarrows C_0$ (note that this pair is reflexive, with common splitting given by s_0). It is clear that π_0 is a functor $\mathbf{Cat}(\mathcal{S}) \to \mathcal{S}$; and it is not hard to verify that it is left adjoint to the functor which sends an object of $\mathcal{S}$ to the corresponding discrete internal category. More interestingly, we have

Lemma 2.5.5 *Let $\mathcal{S}$ and $\mathbb{C}$ be as above. Then the left adjoint $\varinjlim_{\mathbb{C}}$ of $\mathbb{C}^* : \mathcal{S} \to [\mathbb{C}, \mathcal{S}]$ may be described in terms of discrete opfibrations as the functor $(f\colon \mathbb{F} \to \mathbb{C}) \mapsto \pi_0\mathbb{F}$.*

Proof By definition, $\mathbb{C}^*I$ is the object $(C_0)^*I$ of $\mathcal{S}/C_0$ equipped with the identity morphism on $(C_1)^*I$ as its structure map. But by 2.5.4(*a*) this corresponds to the discrete opfibration $\pi_2\colon I \times \mathbb{C} \to \mathbb{C}$; that is, $\mathbb{C}^*$ may be identified with the composite of the 'discrete' embedding $\mathcal{S} \to \mathbf{Cat}(\mathcal{S})$ and the pullback functor $\mathbb{C}^* : \mathbf{Cat}(\mathcal{S}) \to \mathbf{Cat}(\mathcal{S})/\mathbb{C}$. Composing the left adjoints π_0 and $\Sigma_{\mathbb{C}}$ of these two functors yields the description above. □

We note that $\pi_0\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \cong C_0$, since we have a split coequalizer diagram

$$C_2 \underset{d_2}{\overset{d_1}{\rightrightarrows}} C_1 \xrightarrow{d_1} C_0$$

whose splittings are provided by the degeneracy maps s_0^0 and s_0^1 omitted in the definition of $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$. More generally, for any object $x\colon I \to C_0$ of $\mathcal{S}/C_0$ we have $\pi_0\mathbb{R}(x) \cong I$, since the above diagram together with its splittings is preserved by the functor $(-) \times_{C_0} I$. Of course, this latter fact could also have been deduced from the description of $\mathbb{R}$ as the left adjoint of the forgetful functor $\mathcal{S}^{\mathbb{C}} \to \mathcal{S}/C_0$; for its composite with the left adjoint of $\mathbb{C}^*$ must yield the left adjoint Σ_{C_0} of C_0^*.

Corollary 2.5.6 *Suppose coequalizers of reflexive pairs are stable under pullback in $\mathcal{S}$. Then, for an internal category $\mathbb{C}$ in $\mathcal{S}$, the functor $\mathbb{C}^*\colon \mathcal{S} \to [\mathbb{C}, \mathcal{S}]$ is full and faithful iff $\mathbb{C}$ is connected, i.e. $\pi_0\mathbb{C} \cong 1$.*

Proof The hypothesis on $\mathcal{S}$ implies that $C_1 \times I \rightrightarrows C_0 \times I \to \pi_0\mathbb{C} \times I$ is a coequalizer for any object I of $\mathcal{S}$, i.e. $\varinjlim_{\mathbb{C}}\mathbb{C}^*I \cong \pi_0\mathbb{C} \times I$. So $\pi_0\mathbb{C} \cong 1$ implies that the counit of the adjunction $(\varinjlim_{\mathbb{C}} \dashv \mathbb{C}^*)$ is an isomorphism; the converse is just the special case $I = 1$ of this isomorphism. □

The following result is the 'internal' analogue of the result that every functor from a small category to **Set** is expressible as a colimit of representable functors.

Lemma 2.5.7 *Let $\mathcal{S}$ be a cartesian category with coequalizers of reflexive pairs which are stable under pullback. Let $\mathbb{C}$ be an internal category in $\mathcal{S}$, and let F be a diagram of shape $\mathbb{C}$ in $\mathbb{S}$. Then there exists a diagram G of shape $\mathbb{F}^{\mathrm{op}}$ in $[\mathbb{C}, \mathcal{S}]$ (where $\gamma\colon \mathbb{F} \to \mathbb{C}$ is the discrete opfibration corresponding to F) such that $\varinjlim_{\mathbb{F}^{\mathrm{op}}}(G) \cong F$, and such that the image of G under the forgetful functor*

$$[\mathbb{F}^{\mathrm{op}}, [\mathbb{C}, \mathcal{S}]] \longrightarrow [\mathbb{C}, \mathcal{S}]^{F_0}$$

is $\mathbb{R}(\gamma_0)$.

Proof In addition to its left $\mathbb{C}$-action, the object $C_1 \times_{C_0} F_0 \cong F_1$ of $\mathcal{S}/C_0 \times F_0$ comes equipped with a right action of $\mathbb{F}$, given by the composition map $d_1^2\colon F_1 \times_{F_0} F_1 \to F_1$; the two actions commute with each other in the appropriate sense (this is just the associativity of composition in $\mathbb{C}$), and so we may regard it as a diagram of shape $\mathbb{F}^{\mathrm{op}}$ in $[\mathbb{C}, \mathcal{S}]$, or equivalently as a diagram of shape $\mathbb{C}$ in $[\mathbb{F}^{\mathrm{op}}, \mathcal{S}]$. (In the notation which we shall introduce in Section B2.7 below, G is the profunctor $\gamma_\bullet\colon \mathbb{F} \nrightarrow \mathbb{C}$.) Now the diagram

$$F_2 \underset{d_1^2}{\overset{d_0^2}{\rightrightarrows}} F_1 \xrightarrow{d_0^1} F_0$$

is a split coequalizer (the splittings being given by s_0^0 and s_1^1) in $\mathcal{S}/C_0$; and the maps in it are all equivariant for the left actions of $\mathbb{C}$ on the three objects. Also,

the hypotheses on $\mathcal{S}$ ensure that the forgetful functor $[\mathbb{C},\mathcal{S}] \to \mathcal{S}/C_0$ creates coequalizers of reflexive pairs; so the diagram remains a coequalizer in $[\mathbb{C},\mathcal{S}]$. But it is precisely this coequalizer which we have to compute, by 2.5.5, in order to find $\varinjlim_{\mathbb{F}^{\mathrm{op}}}(G)$. □

Classically, it is well known that the functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is the 'free cocompletion' of a small category $\mathcal{C}$. The next result is the analogue for internal categories.

Corollary 2.5.8 *Let $\mathcal{S}$ be as in* 2.5.7, *let $\mathbb{C}$ be an internal category in $\mathcal{S}$ and let $\mathbb{D}$ be a cocomplete $\mathcal{S}$-indexed category. Then the category $[\mathbb{C},\mathbb{D}]$ is equivalent to the category of cocontinuous $\mathcal{S}$-indexed functors $[\mathbb{C}^{\mathrm{op}},\mathbb{S}] \to \mathbb{D}$.*

Proof The object $(d_0,d_1)\colon C_1 \to C_0 \times C_0$ of $\mathcal{S}/C_0 \times C_0$, equipped with both left and right actions of $\mathbb{C}$ in the obvious way, may be regarded as a diagram of shape $\mathbb{C}$ in $[\mathbb{C}^{\mathrm{op}},\mathcal{S}]$. So, given any indexed functor $F\colon [\mathbb{C}^{\mathrm{op}},\mathbb{S}] \to \mathbb{D}$, we obtain a diagram of shape $\mathbb{C}$ in $\mathbb{D}$ by applying $[\mathbb{C},F]$ to this diagram. Conversely, suppose given a diagram G of shape $\mathbb{C}$ in $\mathbb{D}$, which we may think of as an indexed functor $g\colon [\mathbb{C}] \to \mathbb{D}$ via the equivalence of 2.3.13. We have to produce a cocontinuous functor $F\colon [\mathbb{C}^{\mathrm{op}},\mathbb{S}] \to \mathbb{D}$ which agrees with g on representables; but since every object E of $[\mathbb{C}^{\mathrm{op}},\mathcal{S}]$ is a colimit of representables by 2.5.7, there is only one possible way to define this functor. Specifically, we let $\gamma\colon \mathbb{E} \to \mathbb{C}$ be the discrete fibration corresponding to E, and then define $F(E)$ to be the object $\varinjlim_{\mathbb{E}}(\gamma^*G)$ of $\mathcal{D}$. The extension of this definition to an $\mathcal{S}$-indexed functor, the verification that it is cocontinuous, and the proof that the two constructions are inverse to each other up to natural isomorphism, are entirely straightforward. □

It will be recalled that in 1.4.16 we gave a construction for freely adjoining $\mathcal{S}$-indexed coproducts to an $\mathcal{S}$-indexed category. If we apply this to a category of the form $[\mathbb{C}]$, we obtain a category whose I-indexed families of objects are simply objects of $\mathcal{S}/I \times C_0$, and whose I-indexed families of morphisms $(x\colon J \to I \times C_0) \to (y\colon K \to I \times C_0)$ are pairs (f,g) where $f\colon J \to K$ is a morphism over I (though not necessarily over $I \times C_0$) and $g\colon J \to I \times C_1$ satisfies $(1 \times d_1)g = x$ and $(1 \times d_0)g = yf$. If we freely adjoin reflexive coequalizers to this category, then we should expect to obtain the free $\mathcal{S}$-cocompletion of $[\mathbb{C}]$, by the remark after A1.2.10; it is not hard to verify directly that this process would yield an indexed category equivalent to $[\mathbb{C}^{\mathrm{op}},\mathbb{S}]$, so giving an alternative proof of 2.5.8.

Lemma 2.5.5 can be generalized to yield a description of the left adjoint $\varinjlim_f$ of $f^*\colon [\mathbb{D},\mathcal{S}] \to [\mathbb{C},\mathcal{S}]$ for an internal functor $f\colon \mathbb{C} \to \mathbb{D}$ in $\mathcal{S}$. First we show

Lemma 2.5.9 *If $\mathcal{S}$ is a cartesian category with reflexive coequalizers which are stable under pullback, then* $\mathbf{doFib}(\mathcal{S})/\mathbb{C}$ *is reflective in* $\mathbf{Cat}(\mathcal{S})/\mathbb{C}$, *for any internal category $\mathbb{C}$ in $\mathcal{S}$.*

Proof Let $f: \mathbb{D} \to \mathbb{C}$ be an internal functor, and form the coequalizer

$$\begin{tikzcd}
C_1 \times_{C_0} D_1 \arrow[r, shift left, "(d_1(1 \times f_1), d_1\pi_2)"] \arrow[r, shift right, "1 \times d_0"'] \arrow[dr, "d_0\pi_1"'] & C_1 \times_{C_0} D_0 \arrow[r, "q"] \arrow[d, "d_0\pi_1"] & E_0 \arrow[dl, "e_0"] \\
& C_0 &
\end{tikzcd}$$

in $\mathcal{S}/C_0$. Note that the parallel pair in this diagram is reflexive, with common splitting given by $1 \times s_0$; in fact we can make the assignment

$$n \mapsto (d_0\pi_1 : C_1 \times_{C_0} D_n \to C_0)$$

into an internal category $C_1 \rtimes \mathbb{D}$ in an obvious way, and so $e_0 : E_0 \to C_0$ is $\pi_0(C_1 \rtimes \mathbb{D})$. We claim first that e_0 can be given the structure of a discrete opfibration over $\mathbb{C}$. For the hypothesis on coequalizers in $\mathcal{S}$ ensures that

$$\begin{tikzcd}
C_1 \times_{C_0} C_1 \times_{C_0} D_1 \arrow[r, shift left] \arrow[r, shift right] \arrow[dr, "\pi_1"'] & C_1 \times_{C_0} C_1 \times_{C_0} D_0 \arrow[r] \arrow[d, "\pi_1"] & C_1 \times_{C_0} E_0 \arrow[dl, "\pi_1"] \\
& C_1 &
\end{tikzcd}$$

is a coequalizer in $\mathcal{S}/C_1$; now the morphisms $d_1 \times 1 : C_1 \times_{C_0} C_1 \times_{C_0} D_1 \to C_1 \times_{C_0} D_1$ and $d_1 \times 1 : C_1 \times_{C_0} C_1 \times_{C_0} D_0 \to C_1 \times_{C_0} D_0$ induce a morphism $\epsilon : C_1 \times_{C_0} E_0 \to E_0$ making

$$\begin{tikzcd}
C_1 \times_{C_0} E_0 \arrow[r, "\epsilon"] \arrow[d, "\pi_1"] & E_0 \arrow[d, "e_0"] \\
C_1 \arrow[r, "d_0"] & C_0
\end{tikzcd}$$

commute. Together with the second projection $C_1 \times_{C_0} E_0 \to E_0$, this morphism forms part of an internal category structure on $(n \mapsto C_n \times_{C_0} E_0)$, which we denote by $\mathbb{E}$, and e_0 becomes the object-map of a discrete opfibration $\mathbb{E} \to \mathbb{C}$. It is clear, too, that

$$(f : \mathbb{D} \to \mathbb{C}) \mapsto (e : \mathbb{E} \to \mathbb{C})$$

is a functor $\mathbf{Cat}(\mathcal{S})/\mathbb{C} \to \mathbf{doFib}(\mathcal{S})/\mathbb{C}$.

Next, we observe that the composite

$$D_0 \cong C_0 \times_{C_0} D_0 \xrightarrow{s_0 \times 1} C_1 \times_{C_0} D_0 \xrightarrow{q} E_0$$

is the object-map r_0 of an internal functor $r \colon \mathbb{D} \to \mathbb{E}$ over $\mathbb{C}$, where r_n is given by

$$D_n \cong D_n \times_{D_0} D_0 \xrightarrow{f_n \times r_0} C_n \times_{C_0} E_0 \cong E_n \ .$$

And r defines a natural transformation from the identity functor on $\mathbf{Cat}(\mathcal{S})/\mathbb{C}$ to the functor just described. Finally, if f itself is a discrete opfibration, then the split coequalizer diagram

$$\begin{tikzcd} D_2 \arrow[r, shift left, "d_1"] \arrow[r, shift right, "d_0"'] & D_1 \arrow[r, "d_0"] & D_0 \end{tikzcd}$$

(whose splittings are given by $s_0 \colon D_0 \to D_1$ and $s_1 \colon D_1 \to D_2$) is identified by the isomorphisms $D_1 \cong C_1 \times_{C_0} D_0$ and $D_2 \cong C_1 \times_{C_0} D_1$ with the coequalizer defining E_0; thus we conclude that $r_0 \colon D_0 \to E_0$ is an isomorphism, and hence that $r \colon \mathbb{D} \to \mathbb{E}$ is an isomorphism. It follows that, for a general $f \colon \mathbb{D} \to \mathbb{C}$, r is universal among morphisms from f to discrete opfibrations over $\mathbb{C}$. □

Now, if $f \colon \mathbb{C} \to \mathbb{D}$ is an internal functor, it is readily seen that $f^* \colon [\mathbb{D}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ may be described in terms of discrete opfibrations as the pullback functor $f^* \colon \mathbf{Cat}(\mathcal{S})/\mathbb{D} \to \mathbf{Cat}(\mathcal{S})/\mathbb{C}$, restricted to the latter. (It is easy to verify directly that a pullback in $\mathbf{Cat}(\mathcal{S})$ of a discrete opfibration is a discrete opfibration.) Thus we have

Corollary 2.5.10 *Under the hypotheses of* 2.5.9, *the left Kan extension functor* $\varinjlim_f \colon [\mathbb{C}, \mathcal{S}] \to [\mathbb{D}, \mathcal{S}]$ *may be described in terms of discrete opfibrations as the composite*

$$\mathbf{doFib}(\mathcal{S})/\mathbb{C} \longrightarrow \mathbf{Cat}(\mathcal{S})/\mathbb{C} \xrightarrow{\Sigma_f} \mathbf{Cat}(\mathcal{S})/\mathbb{D} \xrightarrow{L} \mathbf{doFib}(\mathcal{S})/\mathbb{D}$$

where L is the reflection functor of 2.5.9. □

Remark 2.5.11 Given two internal functors $f \colon \mathbb{C} \to \mathbb{D}$ and $g \colon \mathbb{B} \to \mathbb{D}$, it is clear that we may construct the comma category $(f \downarrow g)$ in $\mathbf{Cat}(\mathcal{S})$ (its object of objects is $B_0 \times_{D_0} D_1 \times_{D_0} C_0$, and its object of morphisms is an appropriate subobject of $B_1 \times C_1$), and prove that it has the universal property of a comma object (cf. 1.1.4(c)) in $\mathfrak{Cat}(\mathcal{S})$. Using the description of left Kan extensions provided by 2.5.10, it is straightforward to verify that, if $\mathcal{S}$ satisfies the hypotheses

of 2.5.9, we get a commutative square

$$\begin{tikzcd} {[\mathbb{C}, \mathcal{S}]} \arrow[r, "\varinjlim_f"] \arrow[d, "p^*"'] & {[\mathbb{D}, \mathcal{S}]} \arrow[d, "g^*"] \\ {[(f \downarrow g), \mathcal{S}]} \arrow[r, "\varinjlim_q"] & {[\mathbb{B}, \mathcal{S}]} \end{tikzcd}$$

where p and q are the projections. In other words, $\mathbb{S}$ *is pointwise* $\mathcal{A}$*-cocomplete in* $\mathfrak{CAT}_{\mathcal{S}}$ in the sense of 1.1.16(*b*), where $\mathcal{A}$ is as before the class of all functors between internal categories in $\mathcal{S}$. If $\mathcal{S}$ is also locally cartesian closed, then by taking right adjoints of the functors in the above diagram we may also conclude that $\mathbb{S}$ is pointwise $\mathcal{A}$-complete.

We say an internal functor $f : \mathbb{D} \to \mathbb{C}$ is *initial* if its reflection in $\mathbf{doFib}(\mathcal{S})/\mathbb{C}$ is the terminal object $1_{\mathbb{C}}$; equivalently, if the diagram

$$C_1 \times_{C_0} D_1 \underset{1 \times d_0}{\overset{(d_1(1 \times f_1), d_1\pi_2)}{\rightrightarrows}} C_1 \times_{C_0} D_0 \xrightarrow{d_0\pi_1} C_0$$

is a coequalizer. It is easily seen from the composition properties of discrete opfibrations (2.5.2) that the unit of the adjunction of 2.5.9 is always initial; in fact, we have shown that an arbitrary internal functor can be factorized uniquely (up to isomorphism) as an initial functor followed by a discrete opfibration. In the case $\mathcal{S} = \mathbf{Set}$, this factorization structure on $\mathbf{Cat}$ is sometimes known as the *comprehensive* factorization (cf. [1139]).

There is also an important connection between final functors and discrete opfibrations (or equivalently, between initial functors and discrete fibrations). Of course, we say $f : \mathbb{D} \to \mathbb{C}$ is *final* if it is initial as a functor $\mathbb{D}^{\mathrm{op}} \to \mathbb{C}^{\mathrm{op}}$; equivalently, if

$$D_1 \times_{C_0} C_1 \underset{d_1 \times 1}{\overset{(d_0\pi_1, d_1(f_1 \times 1))}{\rightrightarrows}} D_0 \times_{C_0} C_1 \xrightarrow{d_1\pi_2} C_0$$

is a coequalizer. (Traditionally, final functors were called 'cofinal functors'; but this use of 'co' is potentially misleading as it has nothing to do with dualization – it is derived from the Latin 'cum' rather than 'contra' – and so it is now generally omitted.)

Proposition 2.5.12 *Suppose $\mathcal{S}$ is cartesian and has reflexive coequalizers which are stable under pullback. Then an internal functor $f : \mathbb{D} \to \mathbb{C}$ in $\mathcal{S}$ is*

final iff the canonical natural transformation in the diagram

$$\begin{tikzcd}
{[\mathbb{C}, \mathcal{S}]} \arrow[rr, "f^*"] \arrow[dr, "\varinjlim_{\mathbb{C}}"'] & {} \arrow[d, phantom, "\leftarrow"] & {[\mathbb{D}, \mathcal{S}]} \arrow[dl, "\varinjlim_{\mathbb{D}}"] \\
& \mathcal{S} &
\end{tikzcd}$$

(the 'mate' of the natural isomorphism $\mathbb{D}^* \cong f^*\mathbb{C}^*$) is an isomorphism.

Proof First suppose the diagram commutes up to isomorphism. Recall from 2.5.4(b) that we have a particular discrete opfibration $d_0 \colon \mathbb{D}\mathrm{ec}_1(\mathbb{C}) \to \mathbb{C}$, and that $\pi_0 \mathbb{D}\mathrm{ec}_1(\mathbb{C}) \cong C_0$. But if we pull back $\mathbb{D}\mathrm{ec}_1(\mathbb{C})$ along f we obtain the discrete opfibration over $\mathbb{D}$ whose domain is the internal category

$$D_1 \times_{C_0} C_1 \underset{d_1 \times 1}{\overset{(d_0\pi_1, d_1(f_1 \times 1))}{\rightrightarrows}} D_0 \times_{C_0} C_1;$$

so the condition implies that the coequalizer of this pair must be $d_1\pi_2 \colon D_0 \times_{C_0} C_1 \to C_0$.

Conversely, suppose f is final. Then an easy diagram-chase shows that, for any discrete opfibration $e \colon \mathbb{E} \to \mathbb{C}$, the functor q in the pullback

$$\begin{tikzcd}
\mathbb{F} \arrow[r, "q"] \arrow[d, "p"] & \mathbb{E} \arrow[d, "e"] \\
\mathbb{D} \arrow[r, "f"] & \mathbb{C}
\end{tikzcd}$$

is also final; for all the objects and morphisms in the coequalizer diagram for q are obtained simply by pulling back those in the diagram for f along $e_0 \colon E_0 \to C_0$. It therefore suffices to show that if f is final then the canonical morphism $\pi_0\mathbb{D} \to \pi_0\mathbb{C}$ induced by f is an isomorphism. Suppose $u \colon D_0 \to A$ coequalizes $(d_0, d_1) \colon D_1 \rightrightarrows D_0$; then $u\pi_1 \colon D_0 \times_{C_0} C_1 \to A$ coequalizes the pair appearing in the definition of finality, and so we obtain a unique $v \colon C_0 \to A$ satisfying $vd_1\pi_2 = u\pi_1$. Moreover, from the commutativity of the diagram

$$\begin{tikzcd}
D_0 \arrow[r, "{(1, s_0f_0)}"] \arrow[dr, "f_0"'] & D_0 \times_{C_0} C_1 \arrow[r, "\pi_1"] \arrow[d, "d_1\pi_2"] & D_0 \arrow[d, "u"] \\
& C_0 \arrow[r, "v"] & A
\end{tikzcd}$$

we see that we necessarily have $vf_0 = u$; hence v coequalizes $(d_0, d_1)\colon C_1 \rightrightarrows C_0$ and so factors uniquely through $C_0 \twoheadrightarrow \pi_0\mathbb{C}$. Conversely, if $w\colon C_0 \to A$ coequalizes d_0 and d_1, then we necessarily have $wd_1\pi_2 = wf_0\pi_1$; thus we may show that the composite $D_0 \to C_0 \twoheadrightarrow \pi_0\mathbb{C}$ has the universal property of a coequalizer of $D_1 \rightrightarrows D_0$. □

We remark that 2.5.6 is a special case of 2.5.12, since $\mathbb{C} \to \mathbf{1}$ is a final functor iff $\mathbb{C}$ is connected.

B2.6 Filtered colimits

In this and the next section, we shall develop two rather specialized tools of internal category theory, both of which will play crucial rôles in the proof of Diaconescu's Theorem in Section B3.2: the calculus of filtered internal colimits, and the theory of internal profunctors. Both of these can be developed in a context considerably more general than that of a topos (which is why they belong here, rather than in Chapter B3): our standing hypothesis for the present section will be that our ambient category $\mathcal{S}$ is regular and has coequalizers of reflexive pairs.

We saw in 2.3.22 that, for any internal functor $f\colon \mathbb{C} \to \mathbb{D}$ in a topos $\mathcal{S}$, the functor $f^*\colon [\mathbb{D}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ is the inverse image of a geometric morphism. On the other hand, when $\mathcal{S} = \mathbf{Set}$ it is well known that there are important special cases when f^* is a direct image functor, i.e. when its left adjoint $\varinjlim_f$ is cartesian (see, for example, A4.1.10). In this section, we shall consider the corresponding question over a general base category $\mathcal{S}$.

One case when f^* is a direct image functor is when f has a right adjoint $g\colon \mathbb{D} \to \mathbb{C}$ in $\mathfrak{Cat}(\mathcal{S})$; for, since the assignment $\mathbb{C} \mapsto \mathcal{S}^{\mathbb{C}}$ is a 2-functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$, it preserves adjunctions, from which we deduce that f^* is *right* adjoint to g^* – equivalently, that g^* is isomorphic to $\varinjlim_f$. In particular, if $\mathbb{C}$ has a terminal object (which may be defined, for an internal category $\mathbb{C}$, as a morphism $t\colon 1 \to C_0$ for which there exists a morphism $u\colon C_0 \to C_1$ satisfying $d_1u = 1_{C_0}$ and making

$$\begin{tikzcd} C_0 \arrow[r] \arrow[d, "u"] & 1 \arrow[d, "t"] \\ C_1 \arrow[r, "d_0"] & C_0 \end{tikzcd}$$

a pullback), then the unique internal functor $\mathbb{C} \to 1$ has a right adjoint, and so $\varinjlim_{\mathbb{C}}$ is cartesian. However, from our experience of the case $\mathcal{S} = \mathbf{Set}$, we know that this sufficient condition is not necessary; in the $\mathbf{Set}$-based case, $\varinjlim_{\mathbb{C}}$ is cartesian iff $\mathbb{C}$ is a filtered category.

Accordingly, our first task in this section is to formulate the notion of filteredness for an internal category in $\mathcal{S}$. It turns out that the internal logic of a

topos is considerably more than we need in order to do this: the axioms defining filteredness all lie within the regular fragment of first-order logic (cf. D1.1.6), and so can be interpreted in any regular category.

We begin by defining a number of objects associated with an internal category $\mathbb{C}$.

Definition 2.6.1

(*a*) $P_{\mathbb{C}}$, the object of pairs of morphisms of $\mathbb{C}$ with common codomain, is defined by the pullback

$$\begin{tikzcd} P_{\mathbb{C}} \arrow[r, "p_1"] \arrow[d, "p_2"] & C_1 \arrow[d, "d_0"] \\ C_1 \arrow[r, "d_0"] & C_0 \end{tikzcd} .$$

(Note that, because of our convention regarding the symbol $\times_{C_0}$ introduced in Section B2.5, we cannot denote $P_{\mathbb{C}}$ by $C_1 \times_{C_0} C_1$.)

(*b*) $Q_{\mathbb{C}}$, the object of pairs with common domain, is defined by the pullback

$$\begin{tikzcd} Q_{\mathbb{C}} \arrow[r, "q_1"] \arrow[d, "q_2"] & C_1 \arrow[d, "d_1"] \\ C_1 \arrow[r, "d_1"] & C_0 \end{tikzcd} .$$

(c) $R_{\mathbb{C}}$, the object of parallel pairs in $\mathbb{C}$, is the intersection of $P_{\mathbb{C}}$ and $Q_{\mathbb{C}}$ in $\mathrm{Sub}(C_1 \times C_1)$, or equivalently the pullback

$$\begin{tikzcd} R_{\mathbb{C}} \arrow[r, "r_1"] \arrow[d, "r_2"] & C_1 \arrow[d, "(d_0, d_1)"] \\ C_1 \arrow[r, "(d_0, d_1)"] & C_0 \times C_0 \end{tikzcd} .$$

(*d*) $S_{\mathbb{C}}$, the object of commutative squares in $\mathbb{C}$, is defined by the pullback

$$\begin{tikzcd} S_{\mathbb{C}} \arrow[r, "s_1"] \arrow[d, "s_2"] & C_2 \arrow[d, "d_1"] \\ C_2 \arrow[r, "d_1"] & C_1 \end{tikzcd} .$$

(e) $T_{\mathbb{C}}$ is defined by the pullback

$$\begin{tikzcd} T_{\mathbb{C}} \arrow[r, "t_1"] \arrow[d, "t_2"] & C_2 \arrow[d, "(d_0, d_1)"] \\ C_2 \arrow[r, "(d_0, d_1)"] & C_1 \times C_1 \end{tikzcd} \quad ;$$

intuitively, $T_{\mathbb{C}}$ is the object of diagrams of type $\mathcal{T}$ in $\mathbb{C}$, where $\mathcal{T}$ is the finite category

$$\begin{tikzcd} \bullet \arrow[r, shift left, "g"] \arrow[r, shift right, "h"'] & \bullet \arrow[r, "f"] & \bullet \end{tikzcd}$$

with $fg = fh$.

We shall normally omit the subscript $\mathbb{C}$ from the names of the above objects when we are working with a single internal category $\mathbb{C}$.

If we now translate the usual axioms for filteredness of a small category ('there exists an object', 'for any pair of objects, there exists a pair of morphisms from them with common codomain', and 'for any parallel pair of morphisms, there exists a morphism coequalizing them') via the standard interpretation of regular logic in a regular category (cf. D1.2.6), we arrive at the first half of the following definition:

Definition 2.6.2

(a) An internal category $\mathbb{C}$ in a regular category $\mathcal{S}$ is called *filtered* if each of the three morphisms $C_0 \to 1$, $(d_1p_1, d_1p_2): P \to C_0 \times C_0$, and $(d_2t_1, d_2t_2): T \to R$ is a cover.

(b) We say $\mathbb{C}$ is *weakly filtered* if the above morphism $T \to R$ and the morphism $(d_2s_1, d_2s_2): S \to Q$ are both covers.

(Note: in this definition we have adopted the custom of naming morphisms into a pullback by the names of morphisms into the product of which the pullback is (canonically) a subobject. For example, (d_2t_1, d_2t_2) is really a morphism $T \to C_1 \times C_1$, but it can easily be shown to factor through $R \rightarrowtail C_1 \times C_1$.)

We remark that if $\mathbb{C}$ is a preorder (i.e. if $(d_0, d_1): C_1 \to C_0 \times C_0$ is monic), then the condition that $T \to R$ should be a cover is redundant, in both halves of Definition 2.6.2: for in this case the diagonal morphisms $C_1 \to R$ and $C_2 \to T$ are isomorphisms, and (d_2t_1, d_2t_2) is identified with the split epimorphism $d_2: C_2 \to C_1$. For preorders, we commonly use the term *directed* as a synonym for filtered.

We also remark that the assertion that an internal category $\mathbb{C}$ is filtered does not imply the filteredness of the fibres of the corresponding indexed category $[\mathbb{C}]$; however, it is equivalent to the assertion that, given any finite diagram D in some fibre $[\mathcal{C}]^I$ of this indexed category, there exists a cover $x: J \twoheadrightarrow I$ and a

cone under the diagram $x^*(D)$ in $[\mathcal{C}]^J$. (Weak filteredness is equivalent to the same assertion restricted to connected finite diagrams.)

In the proofs of most of the following results, we shall argue informally in terms of elements; that is, we shall prove the assertions in regular logic, and rely on the soundness of the standard interpretation of this logic in a regular category (D1.3.2) to deduce that they hold in $\mathcal{S}$.

Lemma 2.6.3 *A filtered category is weakly filtered.*

Proof Suppose given elements x, y of C_1 satisfying $d_1x = d_1y$. By the second condition in the definition of filteredness, we can find u, v such that $d_0u = d_0v$, $d_1u = d_0x$ and $d_1v = d_0y$. Now the pairs (u, x) and (v, y) belong to C_2, and the pair $(d_1(u, x), d_1(v, y))$ belongs to R, so we can find w with $d_1w = d_0u = d_0v$ and

$$d_1(w, d_1(u, x)) = d_1(w, d_1(v, y)) \ .$$

But the latter equation is equivalent to $d_1(d_1(w, u), x) = d_1(d_1(w, v), y)$; so the elements $(d_1(w, u), x)$ and $(d_1(w, v), y)$ of C_2 define an element of S mapping onto the given element (x, y) of Q. □

Lemma 2.6.4 *If $\mathbb{C}$ is filtered, then it is connected (i.e. $\pi_0\mathbb{C} \cong 1$).*

Proof We know that $\pi_0\mathbb{C} \to 1$ is a cover, since $C_0 \twoheadrightarrow 1$ factors through it. But if x, y are any two elements of C_0, we can find u, v in C_1 with $d_1u = x$, $d_1v = y$ and $d_0u = d_0v$; so if $q\colon C_0 \to \pi_0\mathbb{C}$ is the quotient map we have

$$qx = qd_1u = qd_0u = qd_0v = qd_1v = qy \ .$$

Thus the kernel-pair of q is the whole of $C_0 \times C_0$, and hence $\pi_0\mathbb{C} \to 1$ is monic. □

Proposition 2.6.5 *In an effective regular category $\mathcal{S}$, an internal category $\mathbb{C}$ is filtered iff it is weakly filtered and connected.*

Proof One direction is immediate from the last two lemmas. Conversely, suppose $\mathbb{C}$ is weakly filtered and connected; then clearly $C_0 \to 1$ is a cover, so it only remains to verify the second condition of 2.6.2(a). For this, let $I \rightarrowtail C_0 \times C_0$ be the image of the given morphism $P \to C_0 \times C_0$; we claim first that I is an equivalence relation on C_0, as defined in A1.3.6. The reflexivity and symmetry of I are clear from the definition of P; and transitivity follows from the weak filteredness of $\mathbb{C}$, since if we are given elements x, y, z, w of C_1 satisfying $d_1x = d_1y$, $d_0y = d_0z$ and $d_1z = d_1w$ (so that the pairs (d_0x, d_0y) and (d_0y, d_0w) are both related by I), then the pair (y, z) belongs to Q and we can find u, v with $d_1u = d_1v$, $d_0u = d_1y$ and $d_0v = d_1z$ (and $d_1(y, u) = d_1(z, v)$, although we don't need this). And now the pair $(d_1(x, u), d_1(w, v))$ belongs to P, so that the pair (d_0x, d_0w) are related by I.

Now let $f\colon C_0 \to J$ be a morphism whose kernel-pair is I. Then f factors through the coequalizer $C_0 \to 1$ of d_0 and d_1, since the pair (d_0, d_1)

factors through $P \rightrightarrows C_0$ (and hence through $I \rightrightarrows C_0$) by the morphism $(s_0 d_0, 1_{C_1})\colon C_1 \to P$; hence I must be the whole of $C_0 \times C_0$, and so $P \to C_0 \times C_0$ is a cover. □

For any internal category $\mathbb{C}$, we observe that the $(n+1)$ morphisms $C_n \to C_0$ are all coequalized by $C_0 \to \pi_0\mathbb{C}$; so we can regard each C_n unambiguously as an object of $\mathcal{S}/\pi_0\mathbb{C}$, and all the structural morphisms of $\mathbb{C}$ are morphisms over $\pi_0\mathbb{C}$. Moreover, since $\Sigma_{\pi_0\mathbb{C}}\colon \mathcal{S}/\pi_0\mathbb{C} \to \mathcal{S}$ preserves and reflects pullbacks, we see that $\mathbb{C}$ itself can be regarded as an internal category in $\mathcal{S}/\pi_0\mathbb{C}$; and since the conditions in 2.6.2(b) involve only pullbacks (rather than products), $\mathbb{C}$ is weakly filtered in $\mathcal{S}$ iff it is weakly filtered in $\mathcal{S}/\pi_0\mathbb{C}$. Thus we may rephrase 2.6.5, as follows:

Corollary 2.6.6 *Let $\mathcal{S}$ be an effective regular category. Then an internal category $\mathbb{C}$ in $\mathcal{S}$ is weakly filtered iff it is filtered as an internal category in $\mathcal{S}/\pi_0\mathbb{C}$.* □

Lemma 2.6.7 *Let $f\colon \mathbb{F} \to \mathbb{C}$ be a discrete opfibration. If $\mathbb{C}$ is weakly filtered, then so is $\mathbb{F}$.*

Proof A straightforward diagram-chase shows that the squares

$$\begin{array}{ccc} S_{\mathbb{F}} & \longrightarrow & Q_{\mathbb{F}} \\ \downarrow & & \downarrow \\ S_{\mathbb{C}} & \longrightarrow & Q_{\mathbb{C}} \end{array} \quad \text{and} \quad \begin{array}{ccc} T_{\mathbb{F}} & \longrightarrow & R_{\mathbb{F}} \\ \downarrow & & \downarrow \\ T_{\mathbb{C}} & \longrightarrow & R_{\mathbb{C}} \end{array}$$

are pullbacks, where the vertical arrows are induced by composition with f. So this is immediate from the definition of a regular category. □

Theorem 2.6.8 *Let $\mathcal{S}$ be an effective regular category with coequalizers of reflexive pairs. Then an internal category $\mathbb{C}$ in $\mathcal{S}$ is filtered iff the functor $\varinjlim_{\mathbb{C}}\colon [\mathbb{C}, \mathcal{S}] \to \mathcal{S}$ is cartesian.*

Proof First suppose $\mathbb{C}$ is filtered. We regard $\varinjlim_{\mathbb{C}}$, as in 2.5.5, as the functor sending a discrete opfibration $f\colon \mathbb{F} \to \mathbb{C}$ to $\pi_0\mathbb{F}$; so 2.6.4 tells us that $\varinjlim_{\mathbb{C}}$ preserves the terminal object. So it suffices to show that $\varinjlim_{\mathbb{C}}$ preserves pullbacks; but in fact we need only show that it preserves binary products, since it will then follow from 2.6.6 and 2.6.7 that it sends pullbacks over $(f\colon \mathbb{F} \to \mathbb{C})$ to products in $\mathcal{S}/\pi_0\mathbb{F}$, i.e. to pullbacks over $\pi_0\mathbb{F}$ in $\mathcal{S}$. Now let $f\colon \mathbb{F} \to \mathbb{C}$ and $g\colon \mathbb{G} \to \mathbb{C}$ be two discrete opfibrations; then the projections $\mathbb{F} \times_{\mathbb{C}} \mathbb{G} \to \mathbb{F}$ and $\mathbb{F} \times_{\mathbb{C}} \mathbb{G} \to \mathbb{G}$ induce a morphism $\phi\colon \pi_0(\mathbb{F} \times_{\mathbb{C}} \mathbb{G}) \to \pi_0\mathbb{F} \times \pi_0\mathbb{G}$, and we must show that this is an isomorphism.

To show that ϕ is monic, it suffices to show that any two elements of $F_0 \times_{C_0} G_0$ having the same image in $\pi_0\mathbb{F} \times \pi_0\mathbb{G}$ also have the same image in $\pi_0(\mathbb{F} \times_{\mathbb{C}} \mathbb{G})$.

Let (x_1, y_1) and (x_2, y_2) be two such elements; then x_1 and x_2 have the same image in $\pi_0\mathbb{F}$. But $\mathbb{F}$ is weakly filtered by 2.6.7, so by the argument in the proof of 2.6.5 the pair (x_1, x_2) lies in the image of $P_{\mathbb{F}} \to F_0 \times F_0$; i.e. we can find elements u_1, u_2 of F_1 satisfying $d_1 u_i = x_i$ $(i = 1, 2)$ and $d_0 u_1 = d_0 u_2$. Similarly, we can find elements v_1, v_2 of G_1 satisfying $d_1 v_i = y_i$ and $d_0 v_1 = d_0 v_2$. Now the elements $f_1 u_1$, $f_1 u_2$, $g_1 v_1$ and $g_1 v_2$ of C_1 form a diagram of shape

$$\begin{tikzcd} \bullet \arrow[r] \arrow[dr] & \bullet \\ \bullet \arrow[r] \arrow[ur] & \bullet \end{tikzcd}$$

in $\mathbb{C}$; so we can find elements w_1, w_2 in C_1 with $d_1 w_1 = d_0 f_1 u_i$, $d_1 w_2 = d_0 g_1 v_i$ and $d_0 w_1 = d_0 w_2$. And by an argument similar to that in the proof of 2.6.3 we can also ensure that $d_1(w_1, f_1 u_1) = d_1(w_2, g_1 v_1)$ and $d_1(w_1, f_1 u_2) = d_1(w_2, g_1 v_2)$ as elements of C_1. Now the pair $(w_1, d_0 u_1)$ is in $C_1 \times_{C_0} F_0$, so we may regard it as an element $\tilde{w}_1$ of F_1, and similarly we may lift w_2 to an element $\tilde{w}_2$ of G_1. And the pairs $(d_1(\tilde{w}_1, u_1), d_1(\tilde{w}_2, v_1))$ and $(d_1(\tilde{w}_1, u_2), d_1(\tilde{w}_2, v_2))$ lie in $F_1 \times_{C_1} G_1$; but their domains are the elements (x_1, y_1) and (x_2, y_2) respectively, and their codomains are equal. Hence (x_1, y_1) and (x_2, y_2) have the same image in $\pi_0(\mathbb{F} \times_{\mathbb{C}} \mathbb{G})$.

To show that ϕ is a cover, it similarly suffices to show that, given any element (x, y) of $F_0 \times G_0$, we can find (x', y') in $F_0 \times_{C_0} G_0$ such that x and x' have the same image in $\pi_0\mathbb{F}$, and y and y' have the same image in $\pi_0\mathbb{G}$. But we can find elements u, v of C_1 such that $d_1 u = f_0 x$, $d_1 v = g_0 y$ and $d_0 u = d_0 v$; and if we then lift u and v, as before, to elements $\tilde{u}$ and $\tilde{v}$ of F_1 and G_1 with domains x and y respectively, the elements $x' = d_0 \tilde{u}$ and $y' = d_0 \tilde{v}$ have the required properties.

For the converse, suppose now that $\varinjlim_{\mathbb{C}}$ is cartesian. Since $\varinjlim_{\mathbb{C}}$ preserves the terminal object, we have $\pi_0\mathbb{C} \cong 1$, and so $C_0 \to 1$ is a cover. For the second condition of 2.6.2(a), consider the discrete opfibration $\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \to \mathbb{C}$ of 2.5.4(b): we observed after 2.5.5 that $\pi_0(\mathbb{D}\mathrm{ec}_1(\mathbb{C})) \cong C_0$, so if $\varinjlim_{\mathbb{C}}$ preserves binary products we have $\pi_0(\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \times_{\mathbb{C}} \mathbb{D}\mathrm{ec}_1(\mathbb{C})) \cong C_0 \times C_0$. But the object of objects of $\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \times_{\mathbb{C}} \mathbb{D}\mathrm{ec}_1(\mathbb{C})$ is easily seen to be the pullback $P_{\mathbb{C}}$; so $P \to C_0 \times C_0$ is a cover. The third condition follows similarly from the fact that $\varinjlim_{\mathbb{C}}$ preserves equalizers: the two possible morphisms $d_0 r_1 = d_0 r_2$ and $d_1 r_1 = d_1 r_2 \colon R_{\mathbb{C}} \to C_0$ give rise to discrete opfibrations $\mathbb{R}(d_i r_1) \to \mathbb{C}$, as in 2.5.4($c$), and the two morphisms $R_{\mathbb{C}} \rightrightarrows C_1$ define a pair of morphisms $\mathbb{R}(d_0 r_1) \rightrightarrows \mathbb{R}(d_1 r_1)$ over $\mathbb{C}$. Applying the functor π_0 to the latter, we obtain the identity morphism $R_{\mathbb{C}} \to R_{\mathbb{C}}$ twice over; but if we form the equalizer of the morphisms in $\mathbf{doFib}(\mathcal{S})$ (equivalently, in $\mathbf{Cat}(\mathcal{S})$), we obtain an internal category whose object of objects is readily seen to be isomorphic to $T_{\mathbb{C}}$. So, if $\varinjlim_{\mathbb{C}}$ preserves equalizers, we have a cover $T_{\mathbb{C}} \twoheadrightarrow R_{\mathbb{C}}$; once again, it is easy to verify that this cover is the one appearing in Definition 2.6.2(a). $\square$

Remark 2.6.9 If $\mathbb{C}$ is a filtered internal category in a topos $\mathcal{S}$, we shall sometimes write $\infty_{\mathbb{C}}$ (or simply ∞, if $\mathbb{C}$ is obvious from the context) for the geometric morphism $\mathcal{S} \to [\mathbb{C}, \mathcal{S}]$ whose direct image is $\mathbb{C}^*$ and whose inverse image is $\varinjlim_{\mathbb{C}}$. (Note that, by 2.6.4 and 2.5.6, this is a morphism in $\mathfrak{Top}/\mathcal{S}$.) The thinking behind this notation is that, for any object c of $\mathbb{C}$ (that is, for any morphism $c\colon 1 \to C_0$), 'evaluation at c' is an inverse image functor from $[\mathbb{C}, \mathcal{S}]$ to $\mathcal{S}$, and we think of $\varinjlim_{\mathbb{C}}$ as 'evaluation at the ideal object at infinity'. Note also that, if we merely assume that $\mathbb{C}$ is weakly filtered, then the same two functors define a pre-geometric morphism $\mathcal{S} \to [\mathbb{C}, \mathcal{S}]$ in the sense of A4.1.13, since the proof of 2.6.8 shows that $\varinjlim_{\mathbb{C}}$ preserves pullbacks in this case. (Equivalently, we may regard this as a geometric morphism $\mathcal{S}/\pi_0\mathbb{C} \to [\mathbb{C}, \mathcal{S}]$; cf. 2.6.6.)

Using the description of $\varinjlim_f$ provided by 2.5.10, it is now an easy matter to extend 2.6.8 to a characterization of those internal functors $f\colon \mathbb{D} \to \mathbb{C}$ for which $\varinjlim_f\colon [\mathbb{D}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ is cartesian.

Proposition 2.6.10 *Let $\mathcal{S}$ be an effective regular category in which coequalizers of reflexive pairs exist and are stable under pullback, and let $f\colon \mathbb{D} \to \mathbb{C}$ be an internal functor in $\mathcal{S}$. Then $\varinjlim_f$ is cartesian iff the internal category $C_1 \rtimes \mathbb{D}$ defined in the proof of 2.5.9 is filtered in $\mathcal{S}/C_0$ (equivalently, iff f is initial and $C_1 \rtimes \mathbb{D}$ is weakly filtered in $\mathcal{S}$).*

Proof The forgetful functor $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}/C_0$ is monadic and hence creates finite limits; so $\varinjlim_f$ is cartesian iff its composite with this forgetful functor is. But it is easily verified that the functor $\mathbf{Cat}(\mathcal{S})/\mathbb{C} \to \mathbf{Cat}(\mathcal{S}/C_0)$ which sends $\mathbb{D} \to \mathbb{C}$ to $C_1 \rtimes \mathbb{D}$ is cartesian and preserves discrete opfibrations; so the above composite may be viewed as the composite

$$\mathbf{doFib}(\mathcal{S})/\mathbb{D} \xrightarrow{C_1 \rtimes (-)} \mathbf{doFib}(\mathcal{S}/C_0)/(C_1 \rtimes \mathbb{D}) \xrightarrow{\varinjlim_{C_1 \rtimes \mathbb{D}}} \mathcal{S}/C_0$$

which is cartesian iff its second factor is. This establishes the first characterization; the second follows at once from it and 2.6.6. $\square$

Before leaving this section, we consider the interplay between filtered categories and final functors, as defined at the end of the last section. First we note that, for functors between weakly filtered categories, the definition of final functor may be somewhat simplified; for we have an explicit description of the equivalence relation which is the kernel-pair of $C_0 \twoheadrightarrow \pi_0\mathbb{C}$, as the image of $P_{\mathbb{C}} \to C_0 \times C_0$ (cf. the proof of 2.6.5). Thus we may show

Lemma 2.6.11 *Suppose $\mathcal{S}$ is an effective regular category, and let $\mathbb{D}$ be a filtered internal category in $\mathcal{S}$. Then an internal functor $f\colon \mathbb{D} \to \mathbb{C}$ is final iff the morphisms $d_1\pi_2\colon D_0 \times_{C_0} C_1 \to C_0$ and $((d_1\pi_1, \pi_2)\pi_1, (d_1\pi_1, \pi_2)\pi_2)\colon T_f \to R_f$ are covers, where R_f is the kernel-pair of $(\pi_1, d_1\pi_2)\colon D_0 \times_{C_0} C_1 \to D_0 \times C_0$ and T_f that of $(\pi_1, d_1(f_1 \times 1))\colon D_1 \times_{C_0} C_1 \to D_1 \times C_1$.*

Proof Note that, when f is the identity functor on $\mathbb{C}$, the definitions of R_f and T_f reduce to those of $R_{\mathbb{C}}$ and $T_{\mathbb{C}}$ given in 2.6.1. Intuitively, the second condition says that, given a parallel pair of morphisms $(x,y)\colon c \rightrightarrows f(d)$ in $\mathbb{C}$, we can find $z\colon d \to d'$ such that $f(z)$ has equal composites with x and y.

Since $\mathbb{D}$ is filtered we deduce, as in the proof of 2.6.5, that the equivalence relation generated by $((d_0\pi_1, d_1(f_1 \times 1)), d_1 \times 1)\colon D_1 \times_{C_0} C_1 \rightrightarrows D_0 \times_{C_0} C_1$ is simply the image of $(d_1 \times 1) \times (d_1 \times 1)\colon S_f \to (D_0 \times_{C_0} C_1) \times (D_0 \times_{C_0} C_1)$, where S_f is the kernel-pair of $(d_0\pi_1, d_1(f_1 \times 1))\colon D_1 \times_{C_0} C_1 \to D_0 \times_{C_0} C_1$. So we have to show that the second condition of the lemma is equivalent to saying that $S_f \to Q_f$ is a cover, where Q_f is the kernel-pair of $d_1\pi_2\colon D_0 \times_{C_0} C_1 \to C_0$. But this may easily be proved by 'elementwise' arguments like those in the proofs of 2.6.3 and 2.6.5. □

If $\mathbb{C}$ is also known to be filtered and f is full, then it is easy to verify that the second condition of 2.6.11 is redundant.

In a similar vein, we may establish

Proposition 2.6.12 *Let $f\colon \mathbb{D} \to \mathbb{C}$ be a final internal functor in an effective regular category $\mathcal{S}$. If $\mathbb{D}$ is filtered, then so is $\mathbb{C}$.*

Proof As a sample of the argument, we shall show how $\mathbb{C}$ inherits the second property of 2.6.2(a) from $\mathbb{D}$; the other cases are similar. Consider the diagram

$$\begin{tikzcd}
W \arrow[r] \arrow[d] & (D_0 \times_{C_0} C_1) \times (D_0 \times_{C_0} C_1) \arrow[r, "d_1\pi_2 \times d_1\pi_2"] \arrow[d, "\pi_1 \times \pi_1"] & C_0 \times C_0 \\
P_{\mathbb{D}} \arrow[r] & D_0 \times D_0 &
\end{tikzcd}$$

where the square is a pullback. The top composite is thus a cover; but it is easy to construct a factorization of it through $P_{\mathbb{C}} \to C_0 \times C_0$. □

Of course, given the result of 2.6.8, we could alternatively have deduced the result of 2.6.12 using 2.5.12.

We conclude this section with an important result which, in conjunction with 2.5.12, tells us that we can in practice usually replace arbitrary filtered colimits (that is, colimits over arbitrary filtered categories) by colimits over directed posets. However, for this result we shall need to assume rather more about our base category $\mathcal{S}$.

Theorem 2.6.13 *Let $\mathcal{S}$ be a topos with a natural number object. For any filtered internal category $\mathbb{C}$ in $\mathcal{S}$, there exists a final functor $\mathbb{D} \to \mathbb{C}$ where $\mathbb{D}$ is a directed poset.*

Proof We shall give the proof explicitly only in the case $\mathcal{S} = \mathbf{Set}$: the techniques needed to 'internalize' this proof in an arbitrary topos with natural number object are to be found in Part D. (In particular, we remark that the notion of finiteness required for the proof is the Kuratowskian one, studied in Section D5.4.)

The objects of $\mathbb{D}$ are taken to be finite subcategories, not of $\mathbb{C}$ itself but of $\mathbb{C} \times \mathbb{N}$, where $\mathbb{N}$ is the ordered set of natural numbers (with the ordering defined in Section A2.5); the reason for using this product will become apparent in the course of the proof. We say a subcategory $\mathbb{F}$ of $\mathbb{C} \times \mathbb{N}$ is *suitable* if

(i) $\mathbb{F}$ is finite (that is, both F_0 and F_1 are finite);

(ii) $\mathbb{F}$ has a terminal object; and

(iii) for any morphism $(\alpha, \beta)\colon (c, n) \to (c', n')$ of $\mathbb{F}$, if β is an identity morphism then so is α.

Note that one effect of the third condition is to ensure that the only isomorphisms in $\mathbb{F}$ are identity maps; in particular, this ensures that the terminal object of $\mathbb{F}$ is unique, and we denote it by $t(\mathbb{F})$. We take $\mathbb{D}$ to be the poset of all suitable subcategories of $\mathbb{C} \times \mathbb{N}$, ordered by inclusion. We have an evident functor $f\colon \mathbb{D} \to \mathbb{C}$, which sends $\mathbb{F}$ to the first component of $t(\mathbb{F})$ and an inclusion $\mathbb{F}_1 \subseteq \mathbb{F}_2$ to the first component of the unique morphism $t(\mathbb{F}_1) \to t(\mathbb{F}_2)$ in $\mathbb{F}_2$. We must show that $\mathbb{D}$ is directed and that f is final.

$\mathbb{D}$ is clearly nonempty: if c is any object of $\mathbb{C}$, then the subcategory whose only object is $(c, 0)$ and whose only morphism is the identity is clearly suitable. Let $\mathbb{F}_1$ and $\mathbb{F}_2$ be suitable subcategories. The subcategory $\mathbb{F}_1 \vee \mathbb{F}_2$ generated by their union is still finite, because of condition (iii) which ensures that we do not have to compose arbitrarily long strings of morphisms in the union; and it also clearly satisfies condition (iii). However, it lacks a terminal object in general. Let us write $t(\mathbb{F}_i) = (c_i, n_i)$; let $n_3 = \max\{n_1, n_2\} + 1$, and choose a pair of morphisms $(c_1 \to c_3, c_2 \to c_3)$ in $\mathbb{C}$ with common codomain. If we simply adjoin the morphisms $(c_i, n_i) \to (c_3, n_3)$ $(i = 1, 2)$ to $\mathbb{F}_1 \vee \mathbb{F}_2$, we still may not have a terminal object, since if an object (c_0, n_0) belongs to both $\mathbb{F}_1$ and $\mathbb{F}_2$ then the composites $c_0 \to c_1 \to c_3$ and $c_0 \to c_2 \to c_3$ may not be equal. But there are only finitely many such pairs; so we can find a single morphism $c_3 \to c_4$ coequalizing all of them. Now if we adjoin the morphisms $(c_1, n_1) \to (c_4, n_3)$ and $(c_2, n_2) \to (c_4, n_3)$ (plus all necessary composites) to $\mathbb{F}_1 \vee \mathbb{F}_2$, we obtain a suitable subcategory $\mathbb{F}_3$ containing both $\mathbb{F}_1$ and $\mathbb{F}_2$, as required.

The proof that f is a final functor is very similar, and we omit it. □

Example 2.6.14 To show that the hypothesis that $\mathcal{S}$ has a natural number object cannot be omitted from 2.6.13, consider the case $\mathcal{S} = \mathbf{Set}_f$. Let $M = \{1, e\}$ be the two-element monoid with $e^2 = e$, considered as a finite category (that is, an internal category in $\mathbf{Set}_f$): it is straightforward to verify that M is filtered. However, any finite directed poset P has a greatest member p_0, from which it is easy to see that no functor $f\colon P \to M$ can be final: if $*$ denotes

the unique object of M, then $(1: * \to f(p_0))$ and $(e: * \to f(p_0))$ necessarily lie in different connected components of the comma category $(* \downarrow f)$.

B2.7 Internal profunctors

In this section we develop the basic properties of the bicategory of internal profunctors in a category $\mathcal{S}$; our standing hypothesis for the section is that $\mathcal{S}$ is cartesian and has coequalizers of reflexive pairs which are stable under pullback. By A1.3.5 and the remarks following it, this implies that $\mathcal{S}$ is regular; so this hypothesis is slightly stronger than that which we assumed in the previous section – although we shall not need to assume that $\mathcal{S}$ is effective regular, as we did in the proof of 2.6.5 above.

Classically, a profunctor from a category $\mathcal{C}$ to a category $\mathcal{D}$ (also called a distributor, or a $\mathcal{D}$-$\mathcal{C}$ bimodule) is a functor $\mathcal{C}^{\mathrm{op}} \times \mathcal{D} \to \mathbf{Set}$, thought of as a 'generalized functor' from $\mathcal{C}$ to $\mathcal{D}$. If we replace the ordinary categories $\mathcal{C}$ and $\mathcal{D}$ by internal categories $\mathbb{C}$ and $\mathbb{D}$ in $\mathcal{S}$, then we may of course model this notion by that of a discrete opfibration over $\mathbb{C}^{\mathrm{op}} \times \mathbb{D}$ (or by a discrete fibration over $\mathbb{C} \times \mathbb{D}^{\mathrm{op}}$); but it turns out to be more profitable to exploit the mixture of variances in the definition, producing the following more symmetrical notion:

Definition 2.7.1 Let $\mathbb{C}$ and $\mathbb{D}$ be internal categories in a cartesian category $\mathcal{S}$. By a *profunctor* $T: \mathbb{C} \nrightarrow \mathbb{D}$, we mean an object of $\mathcal{S}/D_0 \times C_0$, or equivalently a span

$$D_0 \longleftarrow T_0 \longrightarrow C_0$$

in $\mathcal{S}$, equipped with a left action $\lambda_T: D_1 \times_{D_0} T_0 \to T_0$ and a right action $\rho_T: T_0 \times_{C_0} C_1 \to T_0$ (both maps over $D_0 \times C_0$), which are associative and unital, and such that

$$\begin{tikzcd} D_1 \times_{D_0} T_0 \times_{C_0} C_1 \arrow[r, "1 \times \rho_T"] \arrow[d, "\lambda_T \times 1"] & D_1 \times_{D_0} T_0 \arrow[d, "\lambda_T"] \\ T_0 \times_{C_0} C_1 \arrow[r, "\rho_T"] & T_0 \end{tikzcd}$$

commutes. A *morphism of profunctors* is a morphism over $D_0 \times C_0$ commuting with the left and right actions. We write $\mathfrak{Prof}(\mathbb{C}, \mathbb{D})$ (or, if necessary, $\mathfrak{Prof}_{\mathcal{S}}(\mathbb{C}, \mathbb{D})$) for the category of profunctors $\mathbb{C} \nrightarrow \mathbb{D}$; by the remarks above, this category is isomorphic to $[\mathbb{C}^{\mathrm{op}} \times \mathbb{D}, \mathcal{S}]$.

We may clearly identify internal diagrams of shape $\mathbb{C}$ and of shape $\mathbb{C}^{\mathrm{op}}$ with profunctors $\mathbf{1} \nrightarrow \mathbb{C}$ and $\mathbb{C} \nrightarrow \mathbf{1}$ respectively, where $\mathbf{1}$ is the terminal object of $\mathbf{Cat}(\mathcal{S})$.

Example 2.7.2 The *unit* or *Yoneda profunctor* $Y(\mathbb{C})\colon \mathbb{C} \looparrowright \mathbb{C}$ is defined to be the object

$$C_0 \xleftarrow{\ d_0\ } C_1 \xrightarrow{\ d_1\ } C_0$$

of $\mathcal{S}/C_0 \times C_0$, with both left and right actions given by $d_1^2\colon C_2 \to C_1$. The condition that the two actions commute with each other reduces in this case to the associativity of composition in $\mathbb{C}$.

We may regard $Y(\mathbb{C})$ as a diagram of shape $\mathbb{C}^{\mathrm{op}}$ in $[\mathbb{C}, \mathcal{S}]$, and hence as an indexed functor $[\mathbb{C}]^{\mathrm{op}} \to [\mathbb{C}, \mathcal{S}]$. As such, it is a full embedding: it sends an object $x\colon I \to C_0$ of $[\mathbb{C}]^I$ to the diagram which we called $\mathbb{R}(x)$ in 2.5.4(c), regarded as an I-indexed family in virtue of the fact (which we observed after 2.5.5) that $\pi_0\mathbb{R}(x) \cong I$. And, using the description of $\mathbb{R}$ as a left adjoint, it is easy to see that morphisms $\mathbb{R}(x) \to \mathbb{R}(y)$ over $\mathbb{C} \times I$ correspond to morphisms $I \to C_1$ making

$$\begin{tikzcd} I \arrow[r] \arrow[dr, "{(x,y)}"'] & C_1 \arrow[d, "{(d_0,d_1)}"] \\ & C_0 \times C_0 \end{tikzcd}$$

commute, i.e. to morphisms $y \to x$ in $[\mathbb{C}]^I$. And since this correspondence is clearly natural in I, it induces an isomorphism

$$[\mathbb{C}]^I(y,x) \cong \mathcal{S}^{\mathbb{C}\times I}(\mathbb{R}(x), \mathbb{R}(y))$$

in $\mathcal{S}/I$.

Now suppose we are given three internal categories $\mathbb{B}, \mathbb{C}, \mathbb{D}$ and two profunctors $S\colon \mathbb{B} \looparrowright \mathbb{C}$, $T\colon \mathbb{C} \looparrowright \mathbb{D}$. We define $(T \otimes_{\mathbb{C}} S)_0$ to be the coequalizer of the pair

$$T_0 \times_{C_0} C_1 \times_{C_0} S_0 \underset{1 \times \lambda_S}{\overset{\rho_T \times 1}{\rightrightarrows}} T_0 \times_{C_0} S_0$$

(note that this pair is reflexive, with common splitting induced by the inclusion-of-identities of $\mathbb{C}$). We note also that $(T \otimes_{\mathbb{C}} S)_0$ comes equipped with a canonical morphism to $D_0 \times B_0$, induced by the morphisms $T_0 \to D_0$ and $S_0 \to B_0$, since the two morphisms above commute with these in the obvious sense. Further, our standing hypothesis on $\mathcal{S}$ implies that the diagram

$$D_1 \times_{D_0} T_0 \times_{C_0} C_1 \times_{C_0} S_0 \underset{1 \times 1 \times \lambda_S}{\overset{1 \times \rho_T \times 1}{\rightrightarrows}} D_1 \times_{D_0} T_0 \times_{C_0} S_0 \longrightarrow D_1 \times_{D_0} (T \otimes_{\mathbb{C}} S)_0$$

is also a coequalizer; and since the morphisms $\rho_T \times 1$ and $1 \times \lambda_S$ are both 'equivariant' in an obvious sense for λ_T, the latter induces a left action

$D_1 \times_{D_0} (T \otimes_{\mathbb{C}} S)_0 \to (T \otimes_{\mathbb{C}} S)_0$. Similarly, ρ_S induces a right action of $\mathbb{B}$ on $(T \otimes_{\mathbb{C}} S)_0$, and it is straightforward to check that these actions give $T \otimes_{\mathbb{C}} S$ the structure of a profunctor $\mathbb{B} \mathrel{\circ\!\!\to} \mathbb{D}$. And it is also clear that if we are given morphisms of profunctors $S \to S'$ and $T \to T'$, we obtain a morphism of profunctors $(T \otimes_{\mathbb{C}} S) \to (T' \otimes_{\mathbb{C}} S')$ in the obvious way; so $(-) \otimes_{\mathbb{C}} (-)$ becomes a bifunctor $\mathfrak{Prof}_S(\mathbb{C},\mathbb{D}) \times \mathfrak{Prof}_S(\mathbb{B},\mathbb{C}) \to \mathfrak{Prof}_S(\mathbb{B},\mathbb{D})$.

Proposition 2.7.3 *Composition of profunctors, as defined above, is associative up to coherent natural isomorphism, and has the Yoneda profunctors $Y(\mathbb{C})$ as two-sided units. Equivalently, profunctors in $\mathcal{S}$ are the* 1*-cells of a bicategory $\mathfrak{Prof}_{\mathcal{S}}$, whose* 0*-cells are the internal categories in $\mathcal{S}$ and whose* 2*-cells are morphisms of profunctors.*

Proof Given profunctors $S\colon \mathbb{B} \mathrel{\circ\!\!\to} \mathbb{C}$, $T\colon \mathbb{C} \mathrel{\circ\!\!\to} \mathbb{D}$ and $U\colon \mathbb{D} \mathrel{\circ\!\!\to} \mathbb{E}$, we may form the diagram

$$\begin{tikzcd}
U_0 \times_{D_0} D_1 \times_{D_0} T_0 \times_{C_0} C_1 \times_{C_0} S_0 \arrow[r, shift left] \arrow[r, shift right] \arrow[d, shift left] \arrow[d, shift right] & U_0 \times_{D_0} T_0 \times_{C_0} C_1 \times_{C_0} S_0 \arrow[r] \arrow[d, shift left] \arrow[d, shift right] & (U \otimes_{\mathbb{D}} T)_0 \times_{C_0} C_1 \times_{C_0} S_0 \arrow[d, shift left] \arrow[d, shift right] \\
U_0 \times_{D_0} D_1 \times_{D_0} T_0 \times_{C_0} S_0 \arrow[r, shift left] \arrow[r, shift right] \arrow[d] & U_0 \times_{D_0} T_0 \times_{C_0} S_0 \arrow[r] \arrow[d] & (U \otimes_{\mathbb{D}} T)_0 \times_{C_0} S_0 \\
U_0 \times_{D_0} D_1 \times_{D_0} (T \otimes_{\mathbb{C}} S)_0 \arrow[r, shift left] \arrow[r, shift right] & U_0 \times_{D_0} (T \otimes_{\mathbb{C}} S)_0 &
\end{tikzcd}$$

in which the first two rows and columns are coequalizers, by our assumption on $\mathcal{S}$; so the coequalizer of the third row is canonically isomorphic to that of the third column. And a straightforward diagram-chase shows that this isomorphism is an isomorphism of profunctors $U \otimes_{\mathbb{D}} (T \otimes_{\mathbb{C}} S) \cong (U \otimes_{\mathbb{D}} T) \otimes_{\mathbb{C}} S$; moreover, the canonical nature of the isomorphism ensures that it satisfies the coherence conditions for composition in a bicategory.

For the units, let us consider $Y(\mathbb{D}) \otimes_{\mathbb{D}} T$, where $T\colon \mathbb{C} \mathrel{\circ\!\!\to} \mathbb{D}$. We have to form the coequalizer of

$$D_1 \times_{D_0} D_1 \times_{D_0} T_0 \underset{1 \times \lambda_T}{\overset{d_1^2 \times 1}{\rightrightarrows}} D_1 \times_{D_0} T_0;$$

but the equations satisfied by $\mathbb{D}$ and T ensure that $\lambda_T\colon D_1 \times_{D_0} T_0 \to T_0$ is a split coequalizer for this pair, with splittings given by

$$D_1 \times_{D_0} D_1 \times_{D_0} T_0 \xleftarrow{s_1^1 \times 1} D_1 \times_{D_0} T_0 \xleftarrow{s_0^0 \times 1} T_0 .$$

Further, λ_T is equivariant for the left action of $\mathbb{D}$ and the right action of $\mathbb{C}$, so it induces an isomorphism of profunctors $Y(\mathbb{D}) \otimes_{\mathbb{D}} T \cong T$. Similarly, we have $T \otimes_{\mathbb{C}} Y(\mathbb{C}) \cong T$; and it is easy to check that these unit isomorphisms satisfy the appropriate coherence conditions with the associativity isomorphisms described above. □

In the case $\mathcal{S} = \mathbf{Set}$, the tensor product of two profunctors $S\colon \mathcal{B}^{\mathrm{op}} \times \mathcal{C} \to \mathbf{Set}$ and $T\colon \mathcal{C}^{\mathrm{op}} \times \mathcal{D} \to \mathbf{Set}$ is often written as a coend:

$$(T \otimes_{\mathcal{C}} S)(B, D) = \int^{C} S(B, C) \times T(C, D)\,.$$

We shall occasionally find it convenient to use this notation for internal profunctors, too.

Now let $f\colon \mathbb{C} \to \mathbb{D}$ be an internal functor. We define a profunctor $f_\bullet\colon \mathbb{C} \nrightarrow \mathbb{D}$ by taking $(f_\bullet)_0$ to be the object $(d_0\pi_1, \pi_2)\colon D_1 \times_{D_0} C_1 \to D_0 \times C_0$ of $\mathcal{S}/D_0 \times C_0$, with right $\mathbb{C}$-action given by

$$D_1 \times_{D_0} C_0 \times_{C_0} C_1 \cong D_1 \times_{D_0} C_1 \xrightarrow{(d_1^2(1 \times f_1),\, d_1^1\pi_2)} D_1 \times_{D_0} C_0$$

and left $\mathbb{D}$-action given by composition in $\mathbb{D}$. Similarly, we may make $(\pi_1, d_1\pi_2)\colon C_0 \times_{D_0} D_1 \to C_0 \times D_0$ into a profunctor $f^\bullet\colon \mathbb{D} \nrightarrow \mathbb{C}$. Note that $(1_{\mathbb{C}})_\bullet$ and $(1_{\mathbb{C}})^\bullet$ are both canonically isomorphic to $Y(\mathbb{C})$.

Proposition 2.7.4

(i) *For any $f\colon \mathbb{C} \to \mathbb{D}$ in $\mathbf{Cat}(\mathcal{S})$, the diagram*

$$\begin{tikzcd}
\mathfrak{Prof}_{\mathcal{S}}(\mathbb{D}, \mathbb{E}) \arrow[r, "(-) \otimes_{\mathbb{D}} (f_\bullet)"] \arrow[d, "\simeq"] & \mathfrak{Prof}_{\mathcal{S}}(\mathbb{C}, \mathbb{E}) \arrow[d, "\simeq"] \\
{[\mathbb{D}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}]} \arrow[r, "(f^{\mathrm{op}} \times 1)^*"] & {[\mathbb{C}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}]}
\end{tikzcd}$$

commutes up to natural isomorphism. (In a similar way, $(f^\bullet) \otimes_{\mathbb{D}} (-)$ may be identified with $(1 \times f)^*\colon [\mathbb{B}^{\mathrm{op}} \times \mathbb{D}, \mathcal{S}] \to [\mathbb{B}^{\mathrm{op}} \times \mathbb{C}, \mathcal{S}]$.)

(ii) *The assignment* $f \mapsto f_\bullet$ *is (the effect on 1-cells of) a pseudofunctor* $\mathfrak{Cat}(\mathcal{S})^{\mathrm{co}} \to \mathfrak{Prof}_{\mathcal{S}}$. *Moreover, this functor is locally full and faithful (i.e. for each pair of internal categories* $(\mathbb{C}, D)$, *the induced functor* $\mathfrak{Cat}(\mathcal{S})(\mathbb{C}, \mathbb{D})^{\mathrm{op}} \to \mathfrak{Prof}_{\mathcal{S}}(\mathbb{C}, \mathbb{D})$ *is full and faithful).*

(iii) *For any* f, *the profunctor* $f^\bullet$ *is right adjoint to* $f_\bullet$ *in* $\mathfrak{Prof}_{\mathcal{S}}$.

Proof (i) This is similar to the verification of the unit law in the proof of 2.7.3. Given a profunctor $T\colon \mathbb{D} \nrightarrow \mathbb{E}$, we have a split coequalizer diagram

$$T_0 \times_{D_0} D_1 \times_{D_0} D_1 \times_{D_0} C_0 \underset{1 \times d_1^2 \times 1}{\overset{\rho_T \times 1 \times 1}{\rightrightarrows}} T_0 \times_{D_0} D_1 \times_{D_0} C_0 \xrightarrow{\rho_T \times 1} T_0 \times_{D_0} C_0$$

with splittings induced by the inclusion-of-identities of $\mathbb{D}$. And it is straightforward to verify that these morphisms are equivariant with respect to the left $\mathbb{E}$-action and the right $\mathbb{C}$-action, when $T_0 \times_{D_0} C_0$ is made into a profunctor $\mathbb{C} \nrightarrow \mathbb{E}$ by regarding it as the pullback of T_0 along $f_0 \times 1$.

(ii) Given internal functors $f\colon \mathbb{C} \to \mathbb{D}$ and $g\colon \mathbb{D} \to \mathbb{E}$, (i) tells us that $(g_\bullet \otimes_{\mathbb{D}} f_\bullet)_0$ is the pullback of $(g_\bullet)_0$ along $(1 \times f_0)$, i.e. it is the object $E_1 \times_{E_0} D_0 \times_{D_0} C_0 \cong E_1 \times_{E_0} C_0$ of $\mathcal{S}/E_0 \times C_0$. As usual, it is straightforward to extend this isomorphism to an isomorphism of profunctors $g_\bullet \otimes_{\mathbb{D}} f_\bullet \cong (gf)_\bullet$. The canonical nature of this isomorphism ensures that it, and the isomorphism $Y(\mathbb{C}) \cong (1_{\mathbb{C}})_\bullet$ which we noted earlier, satisfy the coherence conditions for a pseudofunctor.

Now suppose we are given a 2-cell $\alpha\colon f \to g$ of $\mathfrak{Cat}(\mathcal{S})$, that is an internal natural transformation between internal functors $\mathbb{C} \rightrightarrows \mathbb{D}$. By definition, α is a morphism $C_0 \to D_1$ such that $d_0^1\alpha = g_0$ and $d_1^1\alpha = f_0$, and so we obtain a morphism

$$D_1 \times_{D_0,g_0} C_0 \xrightarrow{1 \times (\alpha,1)} D_1 \times_{D_0} D_1 \times_{D_0,f_0} C_0 \xrightarrow{d_1^2 \times 1} D_1 \times_{D_0,f_0} C_0 ,$$

where we have added extra subscripts to the product signs to indicate the morphisms by which C_0 is regarded as an object over D_0. It is clear that this morphism is equivariant for the left action of $\mathbb{D}$, and (using the naturality of α) for the right action of $\mathbb{C}$: that is, it is a morphism of profunctors $\alpha_\bullet\colon g_\bullet \to f_\bullet$. It is further straightforward to verify that $\alpha \mapsto \alpha_\bullet$ is a functor $\mathfrak{Cat}(\mathcal{S})(\mathbb{C}, \mathbb{D})^{\mathrm{op}} \to \mathfrak{Prof}_{\mathcal{S}}(\mathbb{C}, \mathbb{D})$; and that these functors satisfy the coherence conditions (as $\mathbb{C}$ and $\mathbb{D}$ vary) needed to make $(-)_\bullet$ into a pseudofunctor.

Conversely, suppose we have a morphism of profunctors $h\colon g_\bullet \to f_\bullet$. Then the composite

$$C_0 \xrightarrow{(s_0 g_0, 1)} D_1 \times_{D_0,g_0} C_0 \xrightarrow{h_0} D_1 \times_{D_0,f_0} C_0 \xrightarrow{\pi_1} D_1$$

is a morphism $\alpha\colon C_0 \to D_1$ satisfying $d_0\alpha = g_0$ and $d_1\alpha = f_0$; and it is not hard to see that the $\mathbb{C}$-equivariance condition on h translates into the condition for α to be an internal natural transformation. Moreover, the $\mathbb{D}$-equivariance condition tells us that h can be recovered from its restriction to the subobject $(s_0g_0, 1)\colon C_0 \rightarrowtail D_1 \times_{D_0,g_0} C_0$ of its domain, and hence that the construction just given is a two-sided inverse for the previous construction of profunctor morphisms from internal natural transformations. So $(-)_\bullet\colon \mathfrak{Cat}(\mathcal{S})(\mathbb{C},\mathbb{D})^{\mathrm{op}} \to \mathfrak{Prof}_{\mathcal{S}}(\mathbb{C},\mathbb{D})$ is full and faithful.

(iii) By (i), $(f^\bullet \otimes_{\mathbb{D}} f_\bullet)_0$ is the object $(\pi_1,\pi_3)\colon C_0 \times_{D_0} D_1 \times_{D_0} C_0 \to C_0 \times C_0$; and it is easy to see that $(d_0, f_1, d_1)\colon C_1 \to C_0 \times_{D_0} D_1 \times_{D_0} C_0$ is a morphism of profunctors $\eta_f\colon Y(\mathbb{C}) \to f^\bullet \otimes_{\mathbb{D}} f_\bullet$. On the other hand, $(f_\bullet \otimes_{\mathbb{C}} f^\bullet)_0$ is the coequalizer of

$$\begin{tikzcd}[column sep=huge] D_1 \times_{D_0} C_1 \times_{D_0} D_1 \arrow[r, shift left, "{(\pi_1, d_0\pi_2, d_1(f_1\pi_2, \pi_3))}"] \arrow[r, shift right, "{(d_1(\pi_1, f_1\pi_2), d_1\pi_2, \pi_3)}"'] & D_1 \times_{D_0} C_0 \times_{C_0} D_1; \end{tikzcd}$$

and the morphism $d_1(\pi_1,\pi_3)\colon D_1 \times_{D_0} C_0 \times_{D_0} D_1 \to D_1$ factors through this coequalizer to yield a morphism of profunctors $\epsilon_f\colon f_\bullet \otimes_{\mathbb{C}} f^\bullet \to Y(\mathbb{D})$. It is again straightforward to verify that these morphisms satisfy the triangular identities for an adjunction $(f_\bullet \dashv f^\bullet)$. $\square$

Corollary 2.7.5 *For any $f\colon \mathbb{C} \to \mathbb{D}$ in* $\mathbf{Cat}(\mathcal{S})$, *the diagram*

$$\begin{tikzcd}[column sep=huge, row sep=large] \mathfrak{Prof}_{\mathcal{S}}(\mathbb{C},\mathbb{E}) \arrow[r, "(-)\otimes_{\mathbb{C}}(f^\bullet)"] \arrow[d, "\simeq"] & \mathfrak{Prof}_{\mathcal{S}}(\mathbb{D},\mathbb{E}) \arrow[d, "\simeq"] \\ {[\mathbb{C}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}]} \arrow[r, "\varinjlim_{(f^{\mathrm{op}}\times 1)}"] & {[\mathbb{D}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}]} \end{tikzcd}$$

commutes up to natural isomorphism.

Proof This is immediate from (i) and (iii) of 2.7.4: since $f^\bullet$ is right adjoint to $f_\bullet$, the functor $(-) \otimes f^\bullet$ is left adjoint to $(-) \otimes f_\bullet$. $\square$

Parts (ii) and (iii) of 2.7.4 together say that the passage from $\mathfrak{Cat}(\mathcal{S})$ to the bicategory $\mathfrak{Prof}_{\mathcal{S}}^{\mathrm{co}}$ is very similar to that from a regular category $\mathcal{C}$ to its (locally ordered) 2-category of relations $\mathbf{Rel}(\mathcal{C})$, which we studied in Section A3.1. (This explains why we have re-used for profunctors some of the notation that we used for categories of relations.) However, there is one respect in which the analogy fails to hold: not every morphism of $\mathfrak{Prof}_{\mathcal{S}}$ which has a right adjoint is necessarily of the form $f_\bullet$.

Example 2.7.6 Let $S = \mathbf{Set}$; let $\mathbb{C}$ be a small category, $C \in \mathrm{ob}\,\mathbb{C}$ and let $e\colon C \to C$ be an idempotent endomorphism. Let $F\colon \mathbb{C} \to \mathbf{Set}$ be the functor $D \mapsto \{f\colon C \to D \mid fe = f\}$, and let $G\colon \mathbb{C}^{\mathrm{op}} \to \mathbf{Set}$ be $B \mapsto \{g\colon B \to C \mid eg = g\}$. If we regard F and G as profunctors $\mathbf{1} \nrightarrow \mathbb{C}$ and $\mathbb{C} \nrightarrow \mathbf{1}$ respectively, where $\mathbf{1}$ is the terminal object of $\mathbf{Cat}$, then it is easy to see that F is left adjoint to G. But F is isomorphic to a profunctor of the form $f_\bullet$ iff it is representable, iff the idempotent e splits in $\mathbb{C}$. So if we take $\mathbb{C}$ to be a category in which not all idempotents split, then we have adjoint pairs of profunctors between $\mathbf{1}$ and $\mathbb{C}$ which are not of the form $(f_\bullet, f^\bullet)$.

In fact it can be shown that this is essentially the only possible counterexample: if $\mathbb{C}$ is a Cauchy-complete category and $T\colon \mathbb{B} \nrightarrow \mathbb{C}$ is a profunctor having a right adjoint, then $T \cong f_\bullet$ for some $f\colon \mathbb{B} \to \mathbb{C}$ (which is of course unique up to natural isomorphism, by 2.7.4(ii)). The same result holds, with an appropriate internal definition of Cauchy-completeness, for internal profunctors in any category S satisfying the standing hypothesis of this section. However, we shall not need this result, and so we omit the details of the proof.

Another feature of $\mathfrak{Prof}_S$ can be viewed as the analogue of the fact that $\mathbf{Rel}(\mathcal{C})$ is a tabular allegory:

Proposition 2.7.7 *For any profunctor $T\colon \mathbb{C} \nrightarrow \mathbb{D}$, there is a span*

$$\mathbb{C} \xleftarrow{\ f\ } \mathbb{T} \xrightarrow{\ g\ } \mathbb{D}$$

in $\mathbf{Cat}(S)$ such that T is isomorphic to the profunctor $g_\bullet \otimes_{\mathbb{T}} f^\bullet$.

Proof We use a 'mixed-variance' version of the construction of a discrete opfibration from a diagram, given in 2.5.1. Specifically, we define an internal category $\mathbb{T}$ with object of objects T_0 and object of morphisms given by the pullback

$$\begin{tikzcd} T_1 \arrow[r, "k"] \arrow[d, "h"'] & D_1 \times_{D_0} T_0 \arrow[d, "\lambda_T"] \\ T_0 \times_{C_0} C_1 \arrow[r, "\rho_T"] & T_0 \end{tikzcd} ,$$

with $d_0, d_1\colon T_1 \rightrightarrows T_0$ given by $\pi_1 h$ and $\pi_2 k$ respectively. The inclusion-of-identities of $\mathbb{T}$ is induced in the obvious way by those of $\mathbb{C}$ and $\mathbb{D}$; to define

the composition, consider the diagram

$$\begin{tikzcd}
T_1 \times_{T_0} T_1 \arrow[r, "m"] \arrow[d, "l"] & D_1 \times_{D_0} T_1 \arrow[r, "\pi_2"] \arrow[d, "1 \times h"] & T_1 \arrow[d, "h"] \\
T_1 \times_{C_0} C_1 \arrow[r, "k \times 1"] \arrow[d, "\pi_1"] & D_1 \times_{D_0} T_0 \times_{C_0} C_1 \arrow[r, "\pi_{23}"'] \arrow[d, "\pi_{12}"] & T_0 \times_{C_0} C_1 \arrow[d, "\pi_1"] \\
T_1 \arrow[r, "k"] & D_1 \times_{D_0} T_0 \arrow[r, "\pi_2"] & T_0
\end{tikzcd}$$

in which all the squares are pullbacks. We have

$$d_1\pi_2 h\pi_1 l = f_0\rho_T h\pi_1 l = f_0\lambda_T k\pi_1 l = f_0\pi_2 k\pi_1 l = f_0\pi_1 h\pi_2 m = d_0\pi_2 h\pi_2 m = d_0\pi_2 l$$

and so the pair $(\pi_2 h\pi_1 l, \pi_2 l)$ maps $T_1 \times_{T_0} T_1$ into $C_1 \times_{C_0} C_1$. Hence we can define a morphism

$$T_1 \times_{T_0} T_1 \xrightarrow{(\pi_1 h\pi_1 l,\, d_1(\pi_2 h\pi_1 l, \pi_2 l))} T_0 \times_{C_0} C_1 \,.$$

Similarly, we have

$$T_1 \times_{T_0} T_1 \xrightarrow{(d_1(\pi_1 m, \pi_1 k\pi_2 m), \pi_2 k\pi_2 m)} D_1 \times_{D_0} T_0;$$

and these two morphisms combine to yield $d_1 \colon T_1 \times_{T_0} T_1 \to T_1$. The unit and associative laws for composition in $\mathbb{T}$ follow straightforwardly from those for $\mathbb{C}$ and $\mathbb{D}$.

The projections $D_0 \leftarrow T_0 \rightarrow C_0$ become the object-maps of the internal functors g and f, which are given on morphisms by $\pi_2 h$ and $\pi_1 k$ respectively. To establish the isomorphism $T \cong g_\bullet \otimes_{\mathbb{T}} f^\bullet$, we have to show that

$$D_1 \times_{D_0} T_1 \times_{C_0} C_1 \underset{(d_1(\pi_1, g_1\pi_2), d_1\pi_2, \pi_3)}{\overset{(\pi_1, d_0\pi_2, d_1(f_1\pi_2, \pi_3))}{\rightrightarrows}} D_1 \times_{D_0} T_0 \times_{C_0} C_1 \xrightarrow{\lambda_T(1 \times \rho_T)} T_0$$

is a coequalizer. For this, note first that $1 \times \rho_T \colon D_1 \times_{D_0} T_0 \times_{C_0} C_1 \to D_1 \times_{D_0} T_0$ is the (split) coequalizer of $1 \times 1 \times d_1$ and $1 \times \rho_T \times 1 \colon D_1 \times_{D_0} T_0 \times_{C_0} C_1 \times_{C_0} C_1 \rightrightarrows D_1 \times_{D_0} T_0 \times_{C_0} C_1$; and this pair factors through the pair in the displayed diagram above by

$$D_1 \times_{D_0} T_0 \times_{C_0} C_1 \times_{C_0} C_1 \xrightarrow{1 \times s \times 1} D_1 \times_{D_0} T_1 \times_{C_0} C_1 \,,$$

where s is the splitting of h obtained by pulling back $(s_0 g_0, 1)\colon T_0 \to D_1 \times_{D_0} T_0$ along ρ_T. Hence the coequalizer (q, say) of the pair in the displayed diagram factors through $1 \times \rho_T$. Similarly, q factors through $\lambda_T \times 1$; so it factors through the diagonal of the square

$$\begin{tikzcd} D_1 \times_{D_0} T_0 \times_{C_0} C_1 \arrow[r, "1 \times \rho_T"] \arrow[d, "\lambda_T \times 1"] & D_1 \times_{D_0} T_0 \arrow[d, "\lambda_T"] \\ T_0 \times_{C_0} C_1 \arrow[r, "\rho_T"] & T_0 \end{tikzcd}$$

since this square is a pushout. But it is easy to verify directly that $\lambda_T(1 \times \rho_T)$ coequalizes the given pair; so it is isomorphic to q. $\square$

Corollary 2.7.8 *Suppose S is a locally cartesian closed category with coequalizers of reflexive pairs. Then the bicategory $\mathfrak{Prof}_S$ is biclosed; that is, for each profunctor $T\colon \mathbb{C} \nrightarrow \mathbb{D}$ in S, the functors $(-) \otimes_{\mathbb{D}} T$ and $T \otimes_{\mathbb{C}} (-)$ have right adjoints.*

Proof By 2.7.4(i), 2.7.5 and 2.7.7, we can factor $(-) \otimes_{\mathbb{D}} T\colon \mathfrak{Prof}(\mathbb{D}, \mathbb{E}) \to \mathfrak{Prof}(\mathbb{C}, \mathbb{E})$, up to equivalence, as

$$[\mathbb{D}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}] \xrightarrow{(g \times 1)^*} [\mathbb{T}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}] \xrightarrow{\varinjlim_{f \times 1}} [\mathbb{C}^{\mathrm{op}} \times \mathbb{E}, \mathcal{S}];$$

but both of these functors have right adjoints (the first by 2.3.20, since $\mathbb{S}$ is $\mathcal{S}$-complete by 1.4.7). The argument for $T \otimes_{\mathbb{C}} (-)$ is similar. $\square$

Remark 2.7.9 We note that if T is a diagram of shape $\mathbb{C}$, regarded as a profunctor $\mathbf{1} \nrightarrow \mathbb{C}$, then the category $\mathbb{T}$ constructed in the proof of 2.7.7 is just the domain of the discrete opfibration $\gamma\colon \mathbb{T} \to \mathbb{C}$ which corresponds to T under the equivalence of 2.5.3. In particular, it follows as in the proof of 2.7.8 that $(-) \otimes_{\mathbb{C}} T\colon [\mathbb{C}^{\mathrm{op}}, \mathcal{S}] \to \mathcal{S}$ may be factored as the composite

$$[\mathbb{C}^{\mathrm{op}}, \mathcal{S}] \xrightarrow{\gamma^*} [\mathbb{T}^{\mathrm{op}}, \mathcal{S}] \xrightarrow{\varinjlim_{\mathbb{T}^{\mathrm{op}}}} \mathcal{S}\,.$$

This observation will be of importance in the proof of Diaconescu's Theorem in Section B3.2.

Suggestion for further reading: Kock & Wraith [649].

B3

TOPOSES OVER A BASE

B3.1 $\mathcal{S}$-toposes as $\mathcal{S}$-indexed categories

We are interested in this chapter in studying the 2-category $\mathfrak{Top}/\mathcal{S}$ of toposes over a fixed topos $\mathcal{S}$, through the medium of indexed category theory. We recall the precise definition of this 2-category from Section A4.1: its objects (which we call *toposes defined over* $\mathcal{S}$, or simply $\mathcal{S}$*-toposes*) are geometric morphisms $p\colon \mathcal{E} \to \mathcal{S}$ with codomain $\mathcal{S}$, its morphisms $(q\colon \mathcal{F} \to \mathcal{S}) \to (p\colon \mathcal{E} \to \mathcal{S})$ are pairs (f, α) where $f\colon \mathcal{F} \to \mathcal{E}$ is a geometric morphism and $\alpha : q \cong pf$ a geometric transformation, and its 2-cells $(f, \alpha) \to (g, \beta)$ are geometric transformations $f \to g$ compatible in the obvious sense with α and β. However, we shall almost invariably abuse notation by suppressing any mention of the 2-isomorphism α when specifying 1-cells of $\mathfrak{Top}/\mathcal{S}$; we shall also tend to suppress the structural morphism p when specifying objects of $\mathfrak{Top}/\mathcal{S}$, and simply write '$\mathcal{E}$ is an $\mathcal{S}$-topos'. (The justification for the latter abuse is contained in Theorem 3.1.2 below.)

The main message we wish to convey is that, to a very large extent, the way in which $\mathfrak{Top}/\mathcal{S}$ behaves is independent of any particular features of $\mathcal{S}$: provided we remain within the context of $\mathcal{S}$-indexed categories and functors, we can pretend that $\mathcal{S}$ is simply the topos of sets. (There are only two significant exceptions to this, namely that we must refrain from using the Law of Excluded Middle (resp. the Axiom of Infinity) in our set-theoretic arguments, unless $\mathcal{S}$ happens to be a Boolean topos (resp. to have a natural number object). However, both of these restrictions will be found quite easy to get used to in practice.)

Of course, if $p\colon \mathcal{E} \to \mathcal{S}$ is an $\mathcal{S}$-topos, we use the inverse image functor p^* to make $\mathcal{E}$ into an $\mathcal{S}$-indexed category, as in 1.2.2(d); that is, we set $\mathcal{E}^I = \mathcal{E}/p^*I$. (In fact we shall tend to simplify our notation, whenever possible, by identifying an object I of $\mathcal{S}$ with its image p^*I in $\mathcal{E}$.) As usual, we write $\mathbb{E}$ for the $\mathcal{S}$-indexed category obtained from $\mathcal{E}$ in this way.

Since p^* is cartesian, it extends to an indexed (cartesian) functor $\mathbb{S} \to \mathbb{E}$, by 1.2.2($d$); we shall again use p^* to denote this indexed functor. Similarly, the right adjoint p_* of p^* extends to an indexed functor $\mathbb{E} \to \mathbb{S}$, which is an indexed right adjoint for p^*, by 1.2.3. So we have an 'indexed geometric morphism' from $\mathbb{E}$ to $\mathbb{S}$. We note also that, by 1.4.8 and 2.2.6, $\mathbb{E}$ is cocomplete and locally small as an $\mathcal{S}$-indexed category. Also, if $f\colon \mathcal{F} \to \mathcal{E}$ is any morphism of $\mathfrak{Top}/\mathcal{S}$, then we may

regard the adjunction $(f^* \dashv f_*)$ as indexed over $\mathcal{E}$, and hence (by change of base along the inverse image functor $p^* : \mathcal{S} \to \mathcal{E}$) as an $\mathcal{S}$-indexed adjunction.

However, we observed in 1.2.4 that the left adjoint of f^*, if it has one, does not automatically extend to an $\mathcal{S}$-indexed functor. (A counterexample may be constructed using the second example in 1.4.4(*b*): note that the functor $1 : \mathbf{1} \to \mathbf{2}$ is a discrete opfibration, and so serves to identify $[\mathbf{1}, \mathbf{Set}]$ with a slice category of $[\mathbf{2}, \mathbf{Set}]$.)

Definition 3.1.1 Extending the terminology which we introduced in A4.1.5, we shall say that a geometric morphism $f : \mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ is *$\mathcal{S}$-essential* if f^* has an $\mathcal{S}$-indexed left adjoint $f_!$. Equivalently, f is $\mathcal{S}$-essential if f^* has an ordinary left adjoint $f_!$, and for each morphism $x : I \to J$ in $\mathcal{S}$ the canonical (Beck–Chevalley) natural transformation in the diagram

$$\begin{tikzcd} \mathcal{F}/q^*J \arrow[r, "(f_!)^J"] \arrow[d, "x^*"'] \arrow[dr, phantom, "\to"] & \mathcal{E}/p^*J \arrow[d, "x^*"] \\ \mathcal{F}/q^*I \arrow[r, "(f_!)^I"] & \mathcal{E}/P^*I \end{tikzcd}$$

(where $(f_!)^I$ denotes the functor which sends an object $(B \to q^*I \cong f^*p^*I)$ to its transpose $(f_!B \to p^*I)$, as in 1.2.4) is an isomorphism.

Examples of $\mathcal{S}$-essential morphisms include the geometric morphisms $[\mathbb{C}, \mathcal{S}] \to [\mathbb{D}, \mathcal{S}]$ induced by internal functors $\mathbb{C} \to \mathbb{D}$ in $\mathcal{S}$, as in 2.3.22; the proof of this is contained in 2.3.21.

In the case when $\mathcal{S}$ is **Set**, we have seen (in A4.1.9) that there is (up to isomorphism) at most one geometric morphism from any topos $\mathcal{E}$ to $\mathcal{S}$, so that 'being defined over $\mathcal{S}$' can be regarded as a property of a topos rather than an extra structure. (In fact we saw that it is equivalent to the property of being locally small and having small copowers.) Our first aim in this section is to establish the corresponding result over an arbitrary base $\mathcal{S}$. By an *$\mathcal{S}$-indexed topos*, we mean an $\mathcal{S}$-indexed category $\mathbb{E}$ such that each fibre $\mathcal{E}^I$ is a topos and each transition functor $\mathcal{E}^J \to \mathcal{E}^I$ is logical. (Clearly, the $\mathcal{S}$-indexed category obtained from an $\mathcal{S}$-topos $p : \mathcal{E} \to \mathcal{S}$ as above is an $\mathcal{S}$-indexed topos, by A2.3.3.) An *$\mathcal{S}$-indexed geometric morphism* between $\mathcal{S}$-indexed toposes is, of course, an $\mathcal{S}$-indexed adjunction $(f^* \dashv f_*)$ such that f^* is cartesian. With the obvious definition of $\mathcal{S}$-indexed geometric transformation, we thus have a 2-category $\mathfrak{Top}_{\mathcal{S}}$ of $\mathcal{S}$-indexed toposes. The following result is due to J. L. Moens [820], though we give a later proof due to M. Jibladze [497].

Theorem 3.1.2 *For an $\mathcal{S}$-indexed topos $\mathbb{E}$, the following are equivalent:*

(i) $\mathbb{E}$ *is $\mathcal{S}$-complete and locally small.*

(ii) $\mathbb{E}$ *is $\mathcal{S}$-cocomplete and locally small.*

(iii) *There exists a geometric morphism* $p: \mathcal{E}^1 \to \mathcal{S}$, *such that* $\mathbb{E}$ *is equivalent to the indexed category* $I \mapsto \mathcal{E}^1/p^*I$.

Moreover, if these conditions are satisfied, then the geometric morphism in (iii) *is unique up to unique* 2*-isomorphism.*

Proof (i) $\Leftrightarrow$ (ii) follows from 1.4.6, since by A2.2.10(ii) a logical functor has a left adjoint iff it has a right adjoint (and logical functors are cartesian and cocartesian). We have already noted that (iii) $\Rightarrow$ (ii) follows from 1.4.8 and 2.2.6; so it suffices to prove that (ii) implies (iii).

For this, we employ A2.3.8, which characterized pullback functors amongst logical functors having left adjoints; we shall show that each of the transition functors $x^* : \mathcal{E}^J \to \mathcal{E}^I$ satisfies this characterization. In the case when x is monic, this is easy, since applying the Beck–Chevalley condition to the pullback square

$$\begin{tikzcd} I \arrow[r, "1"] \arrow[d, "1"'] & I \arrow[d, "x"] \\ I \arrow[r, "x"] & J \end{tikzcd}$$

yields the result that the unit $1 \to x^* \circ \Sigma_x$ is an isomorphism, and so Σ_x is (full and) faithful. For a general x, form the pullback square

$$\begin{tikzcd} K \arrow[r, "y"] \arrow[d, "z"'] & I \arrow[d, "x"] \\ I \arrow[r, "x"] & J \end{tikzcd}$$

and the diagonal $w: I \to K$. Given an object A of $\mathcal{E}^I$, we have a commutative diagram

$$\begin{tikzcd} \Sigma_y \Sigma_w w^* z^* A \arrow[r, "\Sigma_y(\epsilon)"] \arrow[d] & \Sigma_y z^* A \arrow[d] \\ A \arrow[r, "\eta"] & x^* \Sigma_x A \end{tikzcd}$$

where η is the unit of $(\Sigma_x \dashv x^*)$ and ϵ is the counit of $(\Sigma_w \dashv w^*)$; the right vertical arrow is an isomorphism by Beck–Chevalley, and the left one is an isomorphism because $yw = zw = 1_I$. Now $\Sigma_w w^*$ is isomorphic to the functor $(-) \times U$, where $U = \Sigma_w(1)$, by A1.5.8; but since w is (split) monic we know from the particular

case already discussed that U must be a subterminal object in $\mathcal{E}^K$. So ϵ is monic; but by A2.4.8 Σ_y preserves monomorphisms, and hence η is monic – equivalently, Σ_x is faithful. So by A2.3.8 we can identify $\mathcal{E}^I$ with a slice of $\mathcal{E}^J$, and x^* with the pullback functor.

In particular, defining p^*I, for $I \in \text{ob}\,\mathcal{S}$, to be Σ_I applied to the terminal object of $\mathcal{E}^I$, we obtain a functor $p^*\colon \mathcal{S} \to \mathcal{E}^1$ such that $\mathbb{E}$ is equivalent to the indexed category $(I \mapsto \mathcal{E}^1/p^*I)$. Thus the latter $\mathcal{S}$-indexed category is cocomplete and locally small, whence p^* is cartesian (since it preserves 1 by definition) and has a right adjoint by 1.4.8 and 2.2.5. So p^* is the inverse image of a geometric morphism, as claimed.

For the uniqueness of p, we observe that it extends to an indexed geometric morphism $\mathbb{E} \to \mathbb{S}$, by the remarks at the beginning of this section; and the latter is unique up to isomorphism by the '$\mathcal{S}$-indexed version' of the argument used in A4.1.9. Specifically, p^* preserves 1 and $\mathcal{S}$-indexed coproducts; but every object $x\colon I \to J$ of $\mathcal{S}/J$ can be written as an $\mathcal{S}$-indexed coproduct of copies of 1 (specifically, $x \cong \Sigma_x(1)$), and so p^* is determined up to unique $\mathcal{S}$-indexed natural isomorphism. Hence the ordinary geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$ is determined up to canonical natural isomorphism (given the indexing of $\mathcal{E}$ over $\mathcal{S}$), as the 1-component of the indexed one. □

Remark 3.1.3 Moens' original proof of the implication (ii)$\Rightarrow$(iii) in 3.1.2 was based on 1.4.12; he showed that if an $\mathcal{S}$-indexed topos has $\mathcal{S}$-indexed coproducts then they are necessarily disjoint and stable (i.e. the '$\mathcal{S}$-indexed analogue' of A2.4.4). We omit the details.

Remark 3.1.4 The assertion, for an $\mathcal{S}$-indexed topos $\mathbb{E}$, that there exists an $\mathcal{S}$-indexed geometric morphism $\mathbb{E} \to \mathbb{S}$, is strictly weaker than the equivalent conditions of 3.1.2. The reason is connected with the fact, noted at the end of A4.1.9, that a topos defined over **Set** need not be cocomplete (in the classical sense); if we take $\mathcal{E}$ to be an incomplete **Set**-topos such as the one described in A2.1.7, then the **Set**-indexing of $\mathcal{E}$ arising from the geometric morphism $\mathcal{E} \to$ **Set** does not coincide with the naive indexing $I \mapsto$ (I-fold power of $\mathcal{E}$) (in fact the latter is the stack completion of the former, for the coverage of 1.5.4); the former is (complete and) cocomplete as a **Set**-indexed category, but the latter is not. However, both indexed categories admit **Set**-indexed geometric morphisms to the canonical indexing of **Set** over itself. On the other hand, it can be shown that an $\mathcal{S}$-indexed topos admitting an indexed geometric morphism to $\mathbb{S}$ is locally small, by the indexed version of the argument of A4.1.9; so we may if we wish add '$\mathbb{E}$ is $\mathcal{S}$-cocomplete and admits an $\mathcal{S}$-indexed geometric morphism to $\mathbb{S}$' to the equivalent conditions of 3.1.2.

Theorem 3.1.2 tells us that we can regard the 2-category $\mathfrak{Top}/\mathcal{S}$ as a sub-2-category of the 2-category $\mathfrak{Top}_{\mathcal{S}}$ of $\mathcal{S}$-indexed toposes. In fact it is a full

sub-2-category, in the appropriate sense:

Proposition 3.1.5 *Let $p\colon \mathcal{E} \to \mathcal{S}$ and $q\colon \mathcal{F} \to \mathcal{S}$ be $\mathcal{S}$-toposes, and let $\mathbb{E}$ and $\mathbb{F}$ be the corresponding $\mathcal{S}$-indexed toposes. Then there is an equivalence of categories*

$$\mathfrak{Top}/\mathcal{S}\,(q,p) \simeq \mathfrak{Top}_{\mathcal{S}}\,(\mathbb{F},\mathbb{E})\;.$$

Proof Given a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$, we extend it to an $\mathcal{S}$-indexed geometric morphism in the obvious way: $(f^*)^I$ is simply f^* applied to objects and morphisms of $\mathcal{E}/p^*I$ (followed by pullback along the isomorphism $q^*I \to f^*p^*I$), and $(f_*)^I$ is f_* followed by pullback along the unit $p^*I \to f_*f^*p^*I \cong f_*q^*I$. Conversely, if $f\colon \mathbb{F} \to \mathbb{E}$ is any $\mathcal{S}$-indexed geometric morphism, then it follows from the uniqueness clause of 3.1.2 that the composite of p with the restriction f^1 of f to the fibres over 1 is isomorphic to q, i.e. that f^1 is a geometric morphism over $\mathcal{S}$. Both constructions are clearly functorial on 2-cells, so it remains to verify that the whole of f is (up to isomorphism) induced by f^1 in the manner indicated above.

Let $g\colon A \to p^*I$ be an object of $\mathcal{E}^I$. Since the square

$$\begin{tikzcd}
\mathcal{E}^I \arrow[r, "(f^*)^I"] \arrow[d, "\Sigma_I"] & \mathcal{F}^I \arrow[d, "\Sigma_I"] \\
\mathcal{E} \arrow[r, "(f^*)^1"] & \mathcal{F}
\end{tikzcd}$$

commutes up to natural isomorphism, we know that $(f^*)^I(g)$ is isomorphic to an object of the form $h\colon (f^*)^1(A) \to q^*I$. And on chasing the graph of g (regarded as a morphism $g \to I^*(A)$ in $\mathcal{E}^I$) around this diagram, and using the fact that f^* also commutes up to isomorphism with I^*, we see that h must be (identified with) $(f^*)^1(g)$, as required. □

We have seen that $\mathcal{S}$-toposes are complete, cocomplete and locally small as $\mathcal{S}$-indexed categories. They are also well-powered and well-copowered, by 1.3.14 and 2.2.7; hence we shall be particularly interested in those which have $\mathcal{S}$-indexed separating and/or coseparating families as defined in 2.4.1, so that we can apply the $\mathcal{S}$-indexed adjoint functor theorem to them. It turns out that separating families for $\mathcal{S}$-toposes can be described in more elementary terms:

Lemma 3.1.6 *Let $p\colon \mathcal{E} \to \mathcal{S}$ be an $\mathcal{S}$-topos, and $\mathbb{E}$ the corresponding $\mathcal{S}$-indexed topos. The following are equivalent:*

(i) *$\mathbb{E}$ has an $\mathcal{S}$-indexed separating family.*

(ii) *There exists an object B of $\mathcal{E}$ such that every object of $\mathcal{E}$ is a subquotient (cf.* A4.6.1*) of one of the form $p^*I \times B$.*

(iii) *There exists an object B of $\mathcal{E}$ such that, for every object A of $\mathcal{E}$, the composite*

$$p^*p_*(\tilde{A}^B) \times B \xrightarrow{\epsilon \times 1} \tilde{A}^B \times B \xrightarrow{\text{ev}} \tilde{A}$$

is epic, where $\tilde{A}$ is the partial-map representer for A (cf. A2.4.7*) and ϵ is the counit of $(p^* \dashv p_*)$.*

Proof (i) ⇒ (ii): Let $g\colon B \to p^*J \in \text{ob}\,\mathcal{E}^J$ be a separating family for $\mathbb{E}$, and let A be any object of $\mathcal{E}$. By 2.4.3, there exists $x\colon I \to J$ in $\mathcal{S}$ and an epimorphism $\Sigma_I x^*(g) \twoheadrightarrow A$ in $\mathcal{E}$; but $\Sigma_I x^*(g)$ is the top left vertex of the pullback square

$$\begin{tikzcd} P \arrow[r] \arrow[d] & B \arrow[d, "g"] \\ p^*I \arrow[r, "p^*x"] & p^*J \end{tikzcd}$$

and hence a subobject of $p^*I \times B$. So the object B has the property specified in (ii).

(ii) ⇒ (iii): Apply (ii) to the object $\tilde{A}$; since all such objects are injective, it must be a quotient of some $p^*I \times B$, say $h\colon p^*I \times B \twoheadrightarrow \tilde{A}$. Transposing successively across the exponential adjunction and $(p^* \dashv p_*)$, we obtain a morphism $\check{h}\colon I \to p_*(\tilde{A}^B)$ such that h is the composite

$$p^*I \times B \xrightarrow{p^*(\check{h}) \times 1} p^*p_*(\tilde{A}^B) \times B \xrightarrow{\epsilon \times 1} \tilde{A}^B \times B \xrightarrow{\text{ev}} \tilde{A};$$

since h is epic, so is $\text{ev}(\epsilon \times 1)$.

(iii) ⇒ (ii): For any A, set $I = p_*(\tilde{A}^B)$; then since A is a subobject of $\tilde{A}$, it is a subquotient of $p^*I \times B$.

(ii) ⇒ (i): Let B be as in (ii). Set $J = p_*(PB)$, and form the pullback

$$\begin{tikzcd} G \arrow[r] \arrow[d, "g"] & \in_B \arrow[d] \\ p^*p_*(PB) \arrow[r, "\epsilon"] & PB \end{tikzcd} \quad ;$$

we claim that $g\colon G \to p^*J$ is an $\mathcal{S}$-indexed separating family for $\mathbb{E}$. Given an object $f\colon A \to p^*I$ of $\mathcal{E}^I$, choose a subquotient representation

$$\begin{tikzcd} S \arrow[r, two heads] \arrow[d, tail] & A \\ p^*K \times B & \end{tikzcd} \quad ;$$

let $s: p^*K \to PB$ be the name of S, regarded as a relation from p^*K to B, and $\check{s}: K \to p_*(PB) = J$ its transpose. The square

$$\begin{array}{ccc} S & \longrightarrow & \in_B \\ \downarrow & & \downarrow \\ p^*K \times B & \xrightarrow{s \times 1} & PB \times B \end{array}$$

is a pullback by the definition of s; but the bottom edge of this square factors through $\epsilon_{PB} \times 1$, so we also have a pullback square

$$\begin{array}{ccc} S & \longrightarrow & G \\ \downarrow & & \downarrow g \\ p^*K & \xrightarrow{p^*(\check{s})} & p^*J \end{array} \quad .$$

Thus, for the span

$$I \xleftarrow{\pi_2} K \times I \xrightarrow{\check{s}\pi_1} J,$$

we obtain an epimorphism $\Sigma_{\pi_2}(\check{s}\pi_1)^*(g) \twoheadrightarrow f$ in $\mathcal{E}^I$, as required. □

Definition 3.1.7 We say an $\mathcal{S}$-topos $p: \mathcal{E} \to \mathcal{S}$ (or the geometric morphism p) is *bounded* if it satisfies the equivalent conditions of Lemma 3.1.6. An object B of $\mathcal{E}$ with the (equivalent) properties described in 3.1.6(ii) or (iii) is called a *bound* for $\mathcal{E}$ over $\mathcal{S}$.

Alternative names that have been used for the concept of bound, as just defined, include 'object of generators' and 'progenitor'.

Examples 3.1.8 (*a*) Localic morphisms, as defined in A4.6.1, are bounded; they are exactly those geometric morphisms such that 1 is a bound for the domain over the codomain. In passing, we note that if a bound B is a subquotient of an object C, then C is also a bound; thus localic morphisms may equivalently be characterized as those for which every well-supported object of the domain is a bound.

(*b*) Interpreting the proof of 3.1.6 in the case $\mathcal{S} = \mathbf{Set}$, we see that a bound for a **Set**-topos $\mathcal{E}$ is simply an object B whose subobjects form a separating set (equivalently, since $\mathcal{E}$ is balanced, a generating set) for $\mathcal{E}$. In particular, if a **Set**-topos $\mathcal{E}$ has a generating set $\{G_i \mid i \in I\}$ such that the coproduct $\coprod_{i \in I} G_i$ exists in $\mathcal{E}$, then the latter is a bound for $\mathcal{E}$, since each G_i occurs as a subobject of it (cf. A2.4.4). For example, in a functor category $[\mathcal{C}, \mathbf{Set}]$

where $\mathcal{C}$ is small, the coproduct of all the representable functors $\mathcal{C}(A, -)$, $A \in$ ob $\mathcal{C}$, is a bound. On the other hand, for a topological group G, the topos **Unif**(G) of uniformly continuous G-sets defined in A2.1.7 need not be bounded over **Set**; it has a generating set (namely the set of discrete quotients of G), but these do not have a coproduct in **Unif**(G) unless G has a smallest open subgroup.

Remark 3.1.9 For an object B of a **Set**-topos $\mathcal{E}$, the assertion that B is a bound for $\mathcal{E}$ over **Set** is easily seen to be equivalent to saying that B is a strong separator for the allegory **Rel**$(\mathcal{E})$, as defined in A3.4.10. Thus one may deduce that *the tabular allegories satisfying the hypotheses of* A3.4.11 *are exactly the categories of relations of bounded* **Set**-*toposes.* Moreover, if $\mathcal{E}$ is a bounded **Set**-topos and B is a bound for $\mathcal{E}$ over **Set**, then we may define an allegory $\mathcal{A}_B$ whose objects are sets and whose morphisms $I \looparrowright J$ are relations in $\mathcal{E}$ from the I-fold copower of B to the J-fold copower; this allegory is positive geometric and has 1 as a strong separator, and if we split its symmetric idempotents we obtain the whole of **Rel**$(\mathcal{E})$ (since every object of $\mathcal{E}$ is a subquotient of a copower of B). So $\mathcal{E}$ may be recovered from $\mathcal{A}_B$ in the manner described in A3.4.12. (However, $\mathcal{A}_B$ will not be pre-tabular unless $B \times B$ happens to be isomorphic to a copower of B; if we want to make it so, we must take its objects to be all set-indexed copowers of finite powers of B.) The point of this observation is that it is often easier to give an explicit construction of the allegory $\mathcal{A}_B$, and to prove that it satisfies the hypotheses of A3.4.12, than to construct $\mathcal{E}$ explicitly and prove directly that it is a topos.

More generally, if $\mathcal{S}$ is any topos whatever, it can be shown that all bounded $\mathcal{S}$-toposes can be constructed from allegories whose objects are those of $\mathcal{S}$, by first splitting symmetric idempotents and then cutting down to the subcategory of maps. To prove this, it is necessary to recast the (set-theoretic) proof of A3.4.11 in terms of $\mathcal{S}$-indexed categories; we shall not go into details here.

Bounded morphisms have properties which are reminiscent of localic morphisms in several ways. For example, given an arbitrary geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$ and a well-supported object B of $\mathcal{E}$, the full subcategory of objects of $\mathcal{E}$ which are subquotients of some $p^*I \times B$ is coreflective – by essentially the same proof as in A4.6.3: to construct the coreflection of an object A, form the image factorization of the morphism obtained by pulling back the morphism displayed in 3.1.6(iii) along $A \rightarrowtail \tilde{A}$ – and is closed under finite limits provided it contains $B \times B$. It is thus a topos, the largest (hyperconnected) quotient of $\mathcal{E}$ for which B is a bound. However, p does not in general have a best possible factorization through a bounded morphism; different choices of B will yield different coreflective subcategories of $\mathcal{E}$, and there may be no largest one.

The next lemma should be compared with A4.6.2(e) and (f), of which it is a direct generalization.

Lemma 3.1.10 *Let*

$$\begin{tikzcd} \mathcal{F} \arrow[r, "f"] \arrow[dr, "q"'] & \mathcal{E} \arrow[d, "p"] \\ & \mathcal{S} \end{tikzcd}$$

be a commutative triangle (up to isomorphism) of geometric morphisms.

(i) *If f and p are both bounded, then q is bounded.*

(ii) *If q is bounded, then f is bounded.*

Proof (i) Since f^* preserves finite products, monomorphisms and epimorphisms, it follows easily from 3.1.6(ii) that if B is a bound for $\mathcal{E}$ over $\mathcal{S}$ and C is a bound for $\mathcal{F}$ over $\mathcal{E}$, then $f^*B \times C$ is a bound for $\mathcal{F}$ over $\mathcal{S}$.

(ii) It is even easier to see that any bound for $\mathcal{F}$ over $\mathcal{S}$ is also a bound for $\mathcal{F}$ over $\mathcal{E}$. □

Corollary 3.1.11 *Let $p\colon \mathcal{E} \to \mathcal{S}$ be a bounded $\mathcal{S}$-topos, and I an object of $\mathcal{S}$. Then the induced geometric morphism $p/I\colon \mathcal{E}/p^*I \to \mathcal{S}/I$ of* A4.1.3 *is also bounded. In fact, if B is a bound for $\mathcal{E}$ over $\mathcal{S}$, then I^*B is a bound for $\mathcal{E}/p^*I$ over $\mathcal{S}/I$.*

Proof Apply both halves of 3.1.10 to the commutative square

$$\begin{tikzcd} \mathcal{E}/p^*I \arrow[r] \arrow[d, "p/I"] & \mathcal{E} \arrow[d, "p"] \\ \mathcal{S}/I \arrow[r] & \mathcal{S} \end{tikzcd}$$

where the horizontal morphisms are localic (A4.6.2(b)) and hence bounded. □

We write $\mathfrak{BTop}/\mathcal{S}$ for the full sub-2-category of $\mathfrak{Top}/\mathcal{S}$ whose objects are bounded $\mathcal{S}$-toposes; thanks to 3.1.10(ii), this is not an abuse of notation. However, in the situation of 3.1.10, the boundedness of the morphism q does not imply that of p, even if f is surjective:

Example 3.1.12 Let $F\colon \mathbf{Set} \to \mathbf{Set}$ be a cartesian functor, and form the topos $\mathbf{Gl}(F)$ as in A2.1.12. $\mathbf{Gl}(F)$ is cocomplete (the forgetful functor $\mathbf{Gl}(F) \to \mathbf{Set} \times \mathbf{Set}$, being comonadic by A4.2.4(c), creates arbitrary colimits) and locally small; so it is defined over **Set** (and its indexing over **Set** is the 'naive' one), and there is a surjection from the (clearly bounded) **Set**-topos $\mathbf{Set} \times \mathbf{Set}$ to it. We claim that $\mathbf{Gl}(F)$ has a generating set iff F is *accessible*, i.e. there is a cardinal κ such that F preserves κ-filtered colimits.

To prove this, first suppose F preserves κ-filtered colimits. Let R be a set containing at least one set of each cardinality less than κ (for example, we could take R to be the power set of a set of cardinality κ), and let G be the set

$$\{(1, 0, 0 \to 1)\} \cup \{(A, 1, \xi) \mid A \in R, \xi : 1 \to F(A)\}$$

where 0 denotes the empty set and 1 a standard singleton set. We claim that G is a generating set for $\mathbf{Gl}(F)$. For if

$$(f, g) : (A, B, \alpha) \longrightarrow (A', B', \alpha')$$

is a monomorphism in $\mathbf{Gl}(F)$ which is not an isomorphism, there are two cases to consider:

(*a*) There exists $a \in A'$ not in the image of f; then we can find a morphism $(1, 0, 0 \to 1) \to (A', B', \alpha')$ not factoring through (f, g).

(*b*) There exists $b \in B'$ not in the image of g; then $\alpha'(b) \in F(A')$ must be in the image of $F(A'') \to F(A')$ for some subset $A'' \subseteq A'$ of cardinality $< \kappa$, and if we choose $\tilde{A} \in R$ of the same cardinality as A'', we then have a morphism $(\tilde{A}, 1, \xi) \to (A', B', \alpha')$ (for a suitable ξ) not factoring through (f, g).

Conversely, suppose $\mathbf{Gl}(F)$ has a generating set G, and let κ be a strict upper bound for the cardinalities of the sets which occur as the first components of members of G. Reversing the above argument, we find that if A is any set and $x \in F(A)$, then in order for us to detect the fact that

$$(1_A, 0 \to 1) : (A, 0, 0 \to F(A)) \longrightarrow (A, 1, x)$$

is not an isomorphism, x must lie in the image of $F(A') \to F(A)$ for some subset $A' \subseteq A$ of cardinality less than κ; from which it follows easily that F preserves κ-filtered colimits.

The question whether every cartesian functor $\mathbf{Set} \to \mathbf{Set}$ is accessible is a nontrivial one, and appears to be independent of the usual axioms of set theory. However, it can be shown (see [121]) that if there is a proper class of measurable cardinals then inaccessible cartesian functors exist. Under this hypothesis, therefore, we have an unbounded $\mathbf{Set}$-topos which is a surjective image of a bounded one.

We conclude this section by considering when the indexed topos $\mathbb{E}$ corresponding to an $\mathcal{S}$-topos $\mathcal{E}$ has a coseparating family. We remark first that if $\mathbb{E}$ has a coseparating family, then it has a single coseparator (that is, a coseparating family living in the fibre over 1); for if $G \in \text{ob } \mathcal{E}^I$ is a coseparating family then so is any object of $\mathcal{E}^I$ containing G as a subobject, in particular the power object $P_I G$; and injectivity of the latter ensures, by a standard argument, that the object $\Pi_I P_I G$ ($\cong P(\Sigma_I G)$) of the fibre $\mathcal{E}^1$ also coseparates $\mathbb{E}$. Further, we have

Proposition 3.1.13 *Let $p: \mathcal{E} \to \mathcal{S}$ be an $\mathcal{S}$-topos. If B is a bound for $\mathcal{E}$ over $\mathcal{S}$, then PB is a coseparator for the corresponding $\mathcal{S}$-indexed topos $\mathbb{E}$.*

Proof Consider first an object A of $\mathcal{E}^1$. By assumption, we have an epimorphism $e: p^*I \times B \twoheadrightarrow \tilde{A}$ for some $I \in \text{ob}\,\mathcal{S}$; and then the composite

$$A \xrightarrow{\ i_A\ } \tilde{A} \xrightarrow{\ \{\}\ } P\tilde{A} \xrightarrow{\ Pe\ } P(p^*I \times B) \cong PB^{p^*I} \cong \Pi_I I^*(PB)$$

is monic by A2.4.7, A2.2.3(i) and A2.3.6(iii). To deal with an object of a general fibre $\mathcal{E}^J$, we argue similarly using 3.1.11. □

The converse of 3.1.13 is false. Let $\mathcal{C}$ be the finite category represented diagrammatically by

$$\bullet \rightrightarrows \bullet;$$

then $[\mathcal{C}, \mathbf{Set}]$ is not localic over $\mathbf{Set}$ (cf. A4.6.9); that is, 1 is not a bound for it. But an easy calculation (cf. [139]) shows that Ω is a coseparator for $[\mathcal{C}, \mathbf{Set}]$. Similarly, the topos $\mathbf{Gl}(F)$ of 3.1.12, though not (in general) bounded over $\mathbf{Set}$, does have a coseparator, namely $(2 \times F(2), 2, \pi_2 : 2 \times F(2) \to F(2))$ where 2 denotes a two-element set; this follows by another easy calculation from the fact that 2 is a coseparator for $\mathbf{Set}$.

Remark 3.1.14 It should be emphasized, before we leave this topic, that the vast majority of geometric morphisms encountered in practice are bounded; in restricting our attention (as we shall do frequently hereafter) to $\mathfrak{BTop}/\mathcal{S}$ rather than $\mathfrak{Top}/\mathcal{S}$, we are not really sacrificing very much. In the case $\mathcal{S} = \mathbf{Set}$, there are essentially only three classes of unbounded $\mathcal{S}$-toposes known to us: those of the form $[\mathcal{G}, \mathbf{Set}]$ where $\mathcal{G}$ is a large groupoid (but with only a set of isomorphism classes of objects, to ensure local smallness of the functor category), or more generally a large category satisfying the condition of A2.1.5; those of the form $\mathbf{Unif}(G)$, where G is a topological group having no smallest open subgroup, as in A2.1.7; and those of the form $\mathbf{Gl}(F)$ where F is an inaccessible cartesian functor $\mathbf{Set} \to \mathbf{Set}$, as in 3.1.12 (or, more generally, those of the form $\mathcal{E}_{\mathbb{G}}$ where $\mathcal{E}$ is a $\mathbf{Set}$-topos and $\mathbb{G}$ is a cartesian comonad on $\mathcal{E}$ whose functor part is inaccessible).

Suggestions for further reading: Adelman & Blass [30], Blass [121], Borceux [139], Jibladze [497], Moens [820].

B3.2 Diaconescu's Theorem

Among the toposes defined over a given base topos $\mathcal{S}$, the diagram categories $[\mathbb{C}, \mathcal{S}]$ defined in Section B2.3 play an important rôle. In this section we investigate them in more detail: our main objective is to prove the important theorem of R. Diaconescu [280] characterizing the geometric morphisms over $\mathcal{S}$ whose codomain is a topos of this form.

First we need the analogue, over an arbitrary base, of a result noted for **Set** in 3.1.8(*b*):

Lemma 3.2.1 *For any* $\mathbb{C}$ *in* $\mathbf{Cat}(\mathcal{S})$, $[\mathbb{C}, \mathcal{S}]$ *is bounded as a topos over* $\mathcal{S}$.

Proof Intuitively, a bound for $[\mathbb{C}, \mathcal{S}]$ over $\mathcal{S}$ should be obtained by taking 'the disjoint union of the representable functors $\mathbb{C} \to \mathcal{S}$': by 2.5.4(c), we can identify the latter as the discrete opfibration $\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \to \mathbb{C}$ of 2.5.4(*b*). To prove that this is indeed a bound, let $f\colon \mathbb{F} \to \mathbb{C}$ be an arbitrary discrete opfibration; then we have a diagram

$$\begin{tikzcd} \mathbb{D}\mathrm{ec}_1(\mathbb{F}) \arrow[r, two heads] \arrow[d, tail] & \mathbb{F} \\ \mathbb{D}\mathrm{ec}_1(\mathbb{C}) \times \mathbb{C}^*(F_0) & \end{tikzcd}$$

of discrete opfibrations over $\mathbb{C}$, where the horizontal arrow is an epimorphism because it is split epic in each dimension, and the vertical arrow is induced in dimension n by the canonical monomorphism

$$F_{n+1} \cong C_{n+1} \times_{C_0} F_0 \rightarrowtail C_{n+1} \times F_0 \, .$$

(Note that these monomorphisms commute with all the face and degeneracy maps of $\mathbb{F}$ and $\mathbb{C}$ except for the face maps d_0 which are dropped in the passage from $\mathbb{F}$ to $\mathbb{D}\mathrm{ec}_1(\mathbb{F})$.) □

Remark 3.2.2 If $\mathbb{C}$ is an internal preorder (that is, if $(d_0, d_1)\colon C_1 \to C_0 \times C_0$ is monic), then we may alternatively take the terminal object of $[\mathbb{C}, \mathcal{S}]$ as a bound, since the canonical morphism $\mathbb{D}\mathrm{ec}_1(\mathbb{C}) \to \mathbb{C} \times C_0$ is monic. Thus, in this case, the geometric morphism $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$ is localic – as we had already noted, in the case $\mathcal{S} = \mathbf{Set}$, in A4.6.2(*d*).

In theory, we have two ways of indexing $[\mathbb{C}, \mathcal{S}]$ over $\mathcal{S}$: that defined in 2.3.12(*a*) from the canonical indexing of $\mathcal{S}$ over itself, and that arising from the geometric morphism $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$. Fortunately, it is not hard to show that these two indexed categories are isomorphic, since we have

$$[\mathbb{C}, \mathcal{S}]/\mathbb{C}^* I \cong \mathbf{doFib}(\mathcal{S})/\mathbb{C} \times I \cong \mathcal{S}^{\mathbb{C} \times I}$$

for any object I of $\mathcal{S}$.

We now introduce the concept of torsor, which will play a fundamental rôle in the main theorem of this section. We recall that if $\mathbb{C}$ is an internal category in a topos $\mathcal{S}$, and $p\colon \mathcal{E} \to \mathcal{S}$ is an $\mathcal{S}$-topos, then diagrams of shape $\mathbb{C}$ in the $\mathcal{S}$-indexed topos $\mathbf{E}$ correspond by 2.3.14 to diagrams of shape $p^*(\mathbb{C})$ in the canonical indexing of $\mathcal{E}$ over itself, and hence to discrete opfibrations over $p^*(\mathbb{C})$ in $\mathbf{cat}(\mathcal{E})$.

Definition 3.2.3 Let $\mathbb{C}$ be an internal category in a topos $\mathcal{S}$, and $p\colon \mathcal{E} \to \mathcal{S}$ an $\mathcal{S}$-topos. By a *torsor* over $\mathbb{C}$ (or $\mathbb{C}$-torsor) in $\mathcal{E}$, we mean an $\mathcal{S}$-indexed functor $F\colon \mathbb{C}^{\mathrm{op}} \to \mathbb{E}$ (that is, a diagram of shape $\mathbb{C}^{\mathrm{op}}$ in $\mathbb{E}$) such that the domain of the corresponding discrete fibration $\mathbb{F} \to p^*(\mathbb{C})$ is filtered (as defined in 2.6.2). We write $\mathbf{Tors}(\mathbb{C}, \mathcal{E})$ for the full subcategory of $[\mathbb{C}^{\mathrm{op}}, \mathcal{E}]$ whose objects are torsors over $\mathbb{C}$.

The term 'torsor' is commonly used for the special case of this concept where the underlying category $\mathbb{C}$ is a group (see 3.2.4(b) below); over a general $\mathbb{C}$, the usual name for the concept is 'flat presheaf'. However, there seems to be no good reason for perpetuating this distinction.

Examples 3.2.4 (a) Let $\mathbb{C}$ be a discrete internal category. Then the domain $\mathbb{F}$ of any discrete fibration $\mathbb{F} \to p^*(\mathbb{C})$ is again discrete (cf. 2.5.4(a)); hence it is filtered iff the unique morphism $F_0 \to 1$ is an isomorphism. Thus $\mathbf{Tors}(\mathbb{C}, \mathcal{E})$ is isomorphic to the discrete category whose objects are the morphisms $1 \to p^*(C_0)$ in $\mathcal{E}$.

(b) Let G be an internal group in $\mathcal{S}$, regarded in the usual way as (the object of morphisms of) an internal groupoid $\mathbb{G}$ with object of objects 1. Then a diagram of shape $\mathbb{G}^{\mathrm{op}}$ in $\mathcal{E}$ is a *right G-object*, i.e. an object F equipped with a unitary, associative action map $\phi\colon F \times p^*G \to F$. Using the fact that the morphisms of $\mathbb{G}$ are all invertible, it is not hard to show that the corresponding total category $\mathbb{F}$ satisfies the second (resp. third) condition of 2.6.2(a) iff the G-action on F is *transitive* (resp. *effective*), i.e. iff the morphism $(\pi_1, \phi)\colon F \times p^*G \to F \times F$ is epic (resp. monic.) Thus a G-torsor is a right G-object F for which $F \to 1$ is epic and (π_1, ϕ) is an isomorphism. Equivalently, it is a G-object F locally isomorphic to p^*G itself: that is, one for which there exists $A \twoheadrightarrow 1$ in $\mathcal{E}$ and an isomorphism $A^*(F) \cong A^*(p^*G)$ of right G-objects in $\mathcal{E}/A$. For if F is a torsor as defined above, then we can take A to be F itself; conversely, if F satisfies the second definition, then $A^*(F \to 1)$ is epic and $A^*(\pi_1, \phi)$ is an isomorphism (since both these assertions are true for $A^*(p^*G)$) and A^* reflects epimorphisms and isomorphisms.

G-torsors arise naturally in cohomology theory, as 'obstructions' to the splitting of epimorphisms; we shall see a good deal more of them in Part E. Note, incidentally, that if F and F' are G-torsors in $\mathcal{E}$, then any G-equivariant map $F \to F'$ is an isomorphism; this may be proved directly from our original definition, but it also follows easily from the second one and the fact that the G-equivariant maps $p^*G \to p^*G$ (the 'left translations' by elements of p^*G) are all isomorphisms. Thus $\mathbf{Tors}(G, \mathcal{E})$ is a groupoid.

(c) For an example of torsors over a monoid which is not a group, we consider the two-element monoid $M = \{1, e\}$ with $e^2 = e$. We may identify M with an internal monoid $\mathbb{M}$ in any topos, by the method of 2.3.12(b); for simplicity, we shall discuss its torsors only in the case $\mathcal{E} = \mathcal{S} = \mathbf{Set}$, but readers who are already familiar with the interpretation of coherent logic in toposes, as described in Section D1.3, will be able to verify that the description which we give is valid

in any topos. An M-set is, of course, just a set A equipped with an idempotent endomorphism $e: A \to A$ (there is no need to distinguish between left and right M-sets, since M is commutative; but we shall write the action on the right, as $a \mapsto a \cdot e$, for consistency); what does it mean to say that (A, e) is an $\mathbb{M}$-torsor?

The first condition of 2.6.2(a) simply says that A is nonempty. The second says that, given any two elements a, b of A, there exists an element c and a pair of morphisms f, g of $\mathbb{M}$ such that $a = c \cdot f$ and $b = c \cdot g$. But there are only four possibilities for the pair (f, g), two of which have $f = g$ and so imply $a = b$. Also, since $ef = e$ for all morphisms f of $\mathbb{M}$, we see that if both a and b satisfy $a = a \cdot e$ and $b = b \cdot e$ then we must have $a = b$; thus there is at most one element in the image of e, and since A is nonempty there must be exactly one such element, a_0 say. Now, for a general pair of elements (a, b), the two possibilities for the pair (f, g) which have $f \neq g$ yield either $a = a_0$ or $b = a_0$. It turns out that the third condition of 2.6.2(a) is automatically satisfied by any M-set; so we deduce that an $\mathbb{M}$-torsor may be identified with a set A equipped with a distinguished element a_0 (the image of the idempotent) and satisfying the coherent axiom

$$(\top \vdash_{x,y} (x = a_0) \vee (y = a_0) \vee (x = y)) .$$

A morphism of $\mathbb{M}$-torsors is simply a map preserving the distinguished elements. As mentioned earlier, the description just given is valid for $\mathbb{M}$-torsors in any topos $\mathcal{E}$; we remark that the displayed axiom implies that the classifying map $A \to \Omega$ of $a_0: 1 \rightarrowtail A$ is monic, and so we may further identify $\mathbb{M}$-torsors in $\mathcal{E}$ with subobjects of Ω which contain the generic subobject $\top$ and satisfy the displayed axiom (with a_0 taken to be $\top$). (We shall meet an important example of an $\mathbb{M}$-torsor in D4.5.12.)

(d) Suppose $\mathbb{C}$ is an internal preorder in $\mathcal{S}$. Then the second condition of 2.6.2(a) implies that, if F is any $\mathbb{C}$-torsor in $\mathcal{E}$, the canonical map $F_0 \to p^*(C_0)$ is monic. Using this, it is easy to see that $\mathbb{C}$-torsors may be identified with what are commonly called *ideals* of $p^*(\mathbb{C})$, that is subobjects of $p^*(C_0)$ which are downwards closed in the ordering and (upwards) directed.

(e) For any internal category $\mathbb{C}$ in $\mathcal{S}$, the (contravariant) décalage $\mathbb{D}\mathrm{ec}^1(\mathbb{C})$ introduced in 2.5.4(b) comes equipped with a discrete fibration $\mathbb{D}\mathrm{ec}^1(\mathbb{C}) \to C_0 \times \mathbb{C}$, which makes it into a $\mathbb{C}$-torsor in $\mathcal{S}/C_0$; for, when regarded as an internal category in $\mathcal{S}/C_0$, $\mathbb{D}\mathrm{ec}^1(\mathbb{C})$ has a terminal object as defined at the beginning of Section B2.6, given by the pullback square

$$\begin{tikzcd} d_0^1 \arrow[r] \arrow[d, "s_0^1"] & 1_{C_0} \arrow[d, "s_0^0"] \\ d_0^1 d_0^2 \arrow[r, "d_0^2"] & d_0^1 \end{tikzcd}$$

in $\mathcal{S}/C_0$.

(*f*) Let $f\colon (\mathcal{F}, q) \to (\mathcal{E}, p)$ be a geometric morphism over $\mathcal{S}$. Since f^* is a regular functor, it preserves discrete fibrations and filtered categories; so it induces a functor $\mathbf{Tors}(\mathbb{C}, \mathcal{E}) \to \mathbf{Tors}(\mathbb{C}, \mathcal{F})$ for any internal category $\mathbb{C}$ in $\mathcal{S}$. There is a converse result if f is a surjection (i.e. f^* is conservative): if $\gamma\colon \mathbb{F} \to p^*(\mathbb{C})$ is any discrete fibration in $\mathcal{E}$ such that $f^*(\gamma)$ corresponds to a $\mathbb{C}$-torsor in $\mathcal{F}$, then γ corresponds to a $\mathbb{C}$-torsor in $\mathcal{E}$. As a particular case of this last observation, we note that the Yoneda profunctor $Y(\mathbb{C})$ of 2.7.2, when regarded as a diagram of shape $\mathbb{C}^{\mathrm{op}}$ in $[\mathbb{C}, \mathcal{S}]$, is always a $\mathbb{C}$-torsor; for its image under the forgetful functor $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}/C_0$ is the torsor corresponding to $\mathbb{D}\mathrm{ec}^1(\mathbb{C})$.

In the case $\mathcal{E} = \mathcal{S} = \mathbf{Set}$, the definition of a torsor may be made more explicit: a functor $F\colon \mathbb{C}^{\mathrm{op}} \to \mathbf{Set}$ is a torsor iff, given any finite diagram $D\colon J \to \mathbb{C}$ and any family of elements $x_j \in F(D(j))$ ($j \in \mathrm{ob}\, J$) which are compatible in the sense that $F(D(\alpha))(x_{j'}) = x_j$ for all $\alpha\colon j \to j'$ in J, there exists a cone $(f_j\colon D(j) \to C \mid j \in \mathrm{ob}\, J)$ under D and an element $x \in F(C)$ such that $F(f_j)(x) = x_j$ for all j. (Of course, it suffices to verify this condition for the special cases when J is the empty category, the discrete category with two objects and the 'parallel-pair' category.) Clearly, in the case when $\mathbb{C}$ is cocartesian, this condition will hold provided F maps finite colimits in $\mathbb{C}$ to limits in $\mathbf{Set}$. But the latter condition is also necessary, and in a more general context than $\mathbf{Set}$:

Lemma 3.2.5 *Let $\mathbb{C}$ be a cocartesian internal category in a topos $\mathcal{S}$ (i.e. one such that the corresponding $\mathcal{S}$-indexed category $[\mathbb{C}]$ is cocartesian), and let $p\colon \mathcal{E} \to \mathcal{S}$ be an $\mathcal{S}$-topos. Then an $\mathcal{S}$-indexed functor $F\colon [\mathbb{C}]^{\mathrm{op}} \to \mathbb{E}$ is a $\mathbb{C}$-torsor iff it is cartesian.*

Proof We observe first that if F is a (contravariant!) diagram of the form $\mathbb{R}(x)$ for a morphism $x\colon 1 \to p^*(C_0)$, then for any object I of $\mathcal{S}$ and any $y\colon I \to C_0 \in \mathrm{ob}\, [\mathbb{C}]^I$, $F^I(y) \cong p_*(J)$, where J is the object of $\mathcal{E}/p^*I$ indexing morphisms $p^*(y) \to x$ in the locally small fibration $[p^*(\mathbb{C})]$, as we defined it in 1.3.12. It is thus immediate from the definition of colimits in $[\mathbb{C}]$ that F transforms them to limits. (Note that diagrams of the form $\mathbb{R}(x)$ are indeed $\mathbb{C}$-torsors, since they are images of $\mathbb{D}\mathrm{ec}^1(\mathbb{C})$ under inverse image functors $\mathcal{S}/C_0 \to \mathcal{E}$.) Now 2.5.7 tells us that a general diagram F of shape $p^*(\mathbb{C}^{\mathrm{op}})$ can be expressed as $\varinjlim_{\mathbb{F}}(G)$, where $\gamma\colon \mathbb{F} \to p^*(\mathbb{C})$ is the discrete fibration corresponding to F and G is a diagram of shape $\mathbb{F}$ in $[p^*(\mathbb{C})^{\mathrm{op}}, \mathcal{E}]$ (equivalently, a diagram of shape $p^*(\mathbb{C})^{\mathrm{op}}$ in $[\mathbb{F}, \mathcal{E}]$) whose underlying diagram in $\mathcal{E}/F_0$ is of the form $\mathbb{R}(\gamma_0)$. It follows that the $\mathcal{S}$-indexed functor corresponding to F can be factored as

$$[\mathbb{C}]^{\mathrm{op}} \xrightarrow{\ G\ } [\mathbb{F}, \mathbb{E}] \xrightarrow{\ \varinjlim_{\mathbb{F}}\ } \mathbb{E};$$

and the first factor is cartesian since the forgetful functor to $\mathbb{E}/F_0$ creates limits, and the second is cartesian by 2.6.8.

For the converse, we first consider the case when $\mathcal{E} = \mathcal{S}$ and p is the identity. Let $\gamma\colon \mathbb{F} \to \mathbb{C}$ be the discrete fibration corresponding to F: by definition, a

morphism $I \to F_0$ corresponds to a morphism $c\colon I \to C_0$ together with a morphism $x\colon 1_I \to F^I(c)$ in $\mathcal{S}/I$. Taking c to be the initial object of $\mathbb{C}$ and x to be the unique morphism $1 \to F^1(c) \cong 1$, we deduce that there exists a morphism $1 \to F_0$, so $F_0 \to 1$ is (split) epic. Next, given two morphisms $I \rightrightarrows F_0$ corresponding to $c, d\colon I \rightrightarrows C_0$ and $x\colon 1_I \to F^I(c)$, $y\colon 1_I \to F^I(d)$, we have a morphism $c + d\colon I \to C_0$ corresponding to the coproduct of c and d in $[\mathbb{C}]^I$ and a morphism $(x, y)\colon 1_I \to F^I(c + d) \cong F^I(c) \times F^I(d)$, yielding a third morphism $I \to F_0$, which is clearly the codomain in $[\mathbb{F}]^I$ of a pair of morphisms whose domains are the two objects of $[\mathbb{F}]^I$ from which we started. Applying this argument to the two projections when $I = F_0 \times F_0$ yields a splitting for the morphism $P_{\mathbb{F}} \to F_0 \times F_0$ of 2.6.2(a). And a similar argument using the existence of coequalizers in $[\mathbb{C}]$ and the fact that F maps them to equalizers will yield a splitting for the morphism $T_{\mathbb{F}} \to R_{\mathbb{F}}$ of 2.6.2(a). So all three of the morphisms in the definition of filteredness for $\mathbb{F}$ are split epic.

Now consider the general case. The assertion '$\mathbb{C}$ is cocartesian' is clearly equivalent to the existence of certain morphisms satisfying certain equations between objects constructed from $\mathbb{C}$ using finite limits (for example, the existence of binary coproducts is equivalent to the existence of a morphism $C_0 \times C_0 \to P_{\mathbb{C}}$ which constructs the coproduct cones, together with a morphism with domain $P_{\mathbb{C}}$ which constructs the factorization of an arbitrary cone through the colimiting one); so it is preserved by cartesian functors, and in particular by the functor p^*. Thus the proof already given, if interpreted in $\mathcal{E}$ rather than in $\mathcal{S}$, shows that F is a $\mathbb{C}$-torsor if the $\mathcal{E}$-indexed functor $F'\colon [p^*(\mathbb{C})]^{\mathrm{op}} \to \mathbb{E}$ which corresponds to F under the equivalence of 2.3.14 is cartesian. At first sight, this condition is stronger than saying 'F is cartesian': the latter is tantamount to saying that F'^J is a cartesian functor for all objects J of $\mathcal{E}$ which happen to be of the form $p^*(I)$, but not for arbitrary objects of $\mathcal{E}$. However, because the 'generic' cocartesian structure of $p^*(\mathbb{C})$ lies in fibres of this kind (for example, any pair of objects of $[p^*(\mathbb{C})]^J$ is obtained by applying an appropriate transition functor of $[p^*(\mathbb{C})]$ to the generic pair (π_1, π_2) in $[p^*(\mathbb{C})]^{p^*(C_0 \times C_0)}$, and an indexed functor which sends the coproduct of this pair to a product will map arbitrary binary coproducts in $[p^*(\mathbb{C})]$ to products), the two conditions (on the $\mathcal{S}$-indexed functor F and the $\mathcal{E}$-indexed functor F') are in fact equivalent. □

Thus we may think of torsors as 'generalized cartesian functors'; informally, they are functors $[\mathbb{C}]^{\mathrm{op}} \to \mathbb{E}$ which 'preserve all finite limits, even when they don't exist in $[\mathbb{C}]^{\mathrm{op}}$'.

Remark 3.2.6 If our base topos $\mathcal{S}$ has a natural number object, then for any internal category $\mathbb{C}$ in $\mathcal{S}$ we may construct the free cocartesian category generated by $\mathbb{C}$, as an internal category. We saw how to construct the free category with finite coproducts generated by $\mathbb{C}$, as an indexed category, in 1.4.17, and in 2.2.4 we saw that this construction preserves local smallness; but it also preserves essential smallness, since if the objects of $\mathbb{C}$ are indexed by C_0 it is easy to see that those of $\mathrm{Fam}_f(\mathbb{C})$ may be indexed by the list object $L(C_0)$ (cf. A2.5.17). Having

thus adjoined finite coproducts to our category, it suffices to give a construction for freely adjoining reflexive coequalizers to a category with finite coproducts; but there is a standard construction (due to A. M. Pitts, cf. [193]) for doing this in the 'external' context, by forming a category whose objects are reflexive pairs in the given category, and it is easy to see that this may be carried out in the internal logic of a topos. Thus we have a functor $\mathbb{C} \mapsto \mathrm{CoCart}(\mathbb{C})$ from $\mathbf{Cat}(\mathcal{S})$ to itself, which sends $\mathbb{C}$ to the free cocartesian category which it generates.

If now $\mathbb{C} \simeq \mathrm{CoCart}(\mathbb{D})$ for some internal category $\mathbb{D}$, then 3.2.5, plus the freeness of the construction, tell us that $\mathbb{C}$-torsors in an $\mathcal{S}$-topos $\mathcal{E}$ correspond to arbitrary $\mathcal{S}$-indexed functors $[\mathbb{D}^{\mathrm{op}}] \to \mathbb{E}$, i.e. to diagrams of shape $\mathbb{D}^{\mathrm{op}}$ in $\mathcal{E}$.

We are now ready to prove the main theorem of this section.

Theorem 3.2.7 (Diaconescu's Theorem) *Let $\mathcal{S}$ be a topos, $\mathbb{C}$ an internal category in $\mathcal{S}$ and $p\colon \mathcal{E} \to \mathcal{S}$ a geometric morphism. Then there is an equivalence of categories*

$$\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{C}, \mathcal{S}]) \simeq \mathbf{Tors}(\mathbb{C}, \mathcal{E}),$$

which is natural in $\mathcal{E}$ in the sense that, if $g\colon \mathcal{F} \to \mathcal{E}$ is a geometric morphism over $\mathcal{S}$, then the diagram

$$\begin{tikzcd}
\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{C}, \mathcal{S}]) \arrow[r, "(-)\circ g"] \arrow[d, "\simeq"] & \mathfrak{Top}/\mathcal{S}\,(\mathcal{F}, [\mathbb{C}, \mathcal{S}]) \arrow[d, "\simeq"] \\
\mathbf{Tors}(\mathbb{C}, \mathcal{E}) \arrow[r, "g^*"] & \mathbf{Tors}(\mathbb{C}, \mathcal{F})
\end{tikzcd}$$

commutes up to (coherent) natural isomorphism.

Proof One direction is immediate from 3.2.4(f): given a geometric morphism $f\colon \mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ over $\mathcal{S}$, we define $\Phi(f)$ to be the $\mathbb{C}$-torsor $f^*(Y(\mathbb{C}))$ in $\mathcal{E}$. It is clear that this yields a functor $\Phi\colon \mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{C}, \mathcal{S}]) \to \mathbf{Tors}(\mathbb{C}, \mathcal{E})$, and that Φ is natural in $\mathcal{E}$ in the sense defined above.

Conversely, suppose given a $\mathbb{C}$-torsor F in $\mathcal{E}$. Consider the composite

$$[\mathbb{C}, \mathcal{S}] \xrightarrow{(p^{\mathbb{C}})^*} [p^*(\mathbb{C}), \mathcal{E}] \xrightarrow{\gamma^*} [\mathbb{F}, \mathcal{E}] \xrightarrow{\varinjlim_{\mathbb{F}}} \mathcal{E},$$

where $(p^{\mathbb{C}})^*$ denotes p^* applied to discrete opfibrations over $\mathbb{C}$, and $\gamma\colon \mathbb{F} \to p^*(\mathbb{C})$ is the discrete fibration corresponding to F. We note that $(p^{\mathbb{C}})^*$ is cartesian and has a right adjoint $(p^{\mathbb{C}})_*$, the latter being obtained by applying p_* to discrete opfibrations over $p^*(\mathbb{C})$ and then pulling back along the unit $\mathbb{C} \to p_*p^*(\mathbb{C})$. So it is the inverse image of a geometric morphism; and so are the other two factors of the composite above, by 2.3.22 and 2.6.9 respectively. Moreover, they are both inverse images of geometric morphisms over $\mathcal{E}$; and since it is easy to verify that

the diagram

$$\begin{tikzcd}
{[p^*(\mathbb{C}), \mathcal{E}]} \arrow[r, "p^{\mathbb{C}}"] \arrow[d] & {[\mathbb{C}, \mathcal{S}]} \arrow[d] \\
\mathcal{E} \arrow[r, "p"] & \mathcal{S}
\end{tikzcd}$$

commutes, it follows that the composite geometric morphism $\Psi(F)$ whose inverse image is displayed above is a morphism over $\mathcal{S}$. We note further that the composite $\varinjlim_{\mathbb{F}}\gamma^*: [p^*(\mathbb{C}), \mathcal{E}] \to \mathcal{E}$ may also be written as $F \otimes_{p^*(\mathbb{C})} (-)$, where we regard F as a profunctor $p^*(\mathbb{C}) \looparrowright \mathbf{1}$ in $\mathcal{E}$, by 2.7.9; from this it follows easily that the construction $F \mapsto \Psi(F)$ is a functor $\mathbf{Tors}(\mathbb{C}, \mathcal{E}) \to \mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{C}, \mathcal{S}])$.

Now, given a torsor F, we have

$$\Phi\Psi(F) = F \otimes_{p^*(\mathbb{C})} (p^{\mathbb{C}})^*(Y(\mathbb{C})) \cong F \otimes Y(p^*(\mathbb{C})) \cong F$$

by 2.7.3. Similarly, given a geometric morphism $f: \mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ over $\mathcal{S}$ and an object G of $[\mathbb{C}, \mathcal{S}]$, we have

$$\begin{aligned}
(\Psi\Phi(f))^*(G) &= f^*(Y(\mathbb{C})) \otimes_{p^*(\mathbb{C})} (p^{\mathbb{C}})^*(G) \\
&\cong f^*(Y(\mathbb{C})) \otimes_{p^*(\mathbb{C})} (f^{\mathbb{C}})^*(q^{\mathbb{C}})^*(G) \\
&\cong f^*(Y(\mathbb{C}) \otimes_{\mathbb{C}} G) \cong f^*(G)
\end{aligned}$$

where we have temporarily written q for the name of the geometric morphism $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$, and used the fact that f^* preserves tensor products of profunctors. Clearly, these isomorphisms are natural in G and in f; hence Ψ is a two-sided inverse for Φ up to natural isomorphism. □

Remarks 3.2.8 (*a*) We recall that any $\mathcal{S}$-topos is $\mathcal{S}$-cocomplete, locally small and well-copowered, and that $[\mathbb{C}, \mathbb{S}]$ also has an $\mathcal{S}$-indexed separating family (3.2.1), so that any $\mathcal{S}$-cocontinuous indexed functor $[\mathbb{C}, \mathbb{S}] \to \mathbb{E}$ has an indexed right adjoint by 2.4.6. Thus, if $\mathbb{C}$ is cocartesian, we may regard Diaconescu's Theorem as expressing the fact that $[\mathbb{C}, \mathbb{S}]$ is the 'free $\mathcal{S}$-cocompletion' of $[\mathbb{C}]^{\mathrm{op}}$, not only in the 2-category $\mathfrak{Cat}_{\mathcal{S}}$ of $\mathcal{S}$-indexed categories (cf. 2.5.8), but also in the 2-category $\mathfrak{Cart}_{\mathcal{S}}$ of cartesian $\mathcal{S}$-indexed categories and functors: that is, composition with the Yoneda embedding $\mathbb{C}^{\mathrm{op}} \to [\mathbb{C}, \mathbb{S}]$ induces an equivalence between cocontinuous cartesian $\mathcal{S}$-indexed functors $[\mathbb{C}, \mathbb{S}] \to \mathbb{E}$ and cartesian functors $\mathbb{C}^{\mathrm{op}} \to \mathbb{E}$. We shall explore this point of view further in Section B4.5 below.

(*b*) Combining 3.2.7 and 3.2.4(*a*), we deduce that for any object I of $\mathcal{S}$ and any $\mathcal{S}$-topos $p: \mathcal{E} \to \mathcal{S}$, $\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, \mathcal{S}/I)$ is equivalent to the discrete category whose objects are the morphisms $1 \to p^*I$ in $\mathcal{E}$. Specializing to the case when $\mathcal{E}$ is itself of the form $\mathcal{S}/J$ for some J, we obtain the information that *the functor* $\mathcal{S} \to \mathfrak{Top}/\mathcal{S}$ *of* A4.1.2, *which sends I to $\mathcal{S}/I$, is a (2-categorical) full embedding.*

(c) On the other hand, the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$ of 2.3.22, which sends $\mathbb{C}$ to $[\mathbb{C}, \mathcal{S}]$, is not full. Using 3.2.7, we may characterize the geometric morphisms

$[\mathbb{C}, \mathcal{S}] \to [\mathbb{D}, \mathcal{S}]$ over $\mathcal{S}$: they correspond to $\mathbb{D}$-torsors in $[\mathbb{C}, \mathcal{S}]$, or equivalently to profunctors $T\colon \mathbb{D} \nrightarrow \mathbb{C}$ which are 'torsors for the right action of $\mathbb{D}$', i.e. such that the functor $T \otimes_{\mathbb{D}} (-)\colon \mathfrak{Prof}_{\mathcal{S}}(\mathbb{E}, \mathbb{D}) \to \mathfrak{Prof}_{\mathcal{S}}(\mathbb{E}, \mathbb{C})$ is cartesian for any $\mathbb{E}$. (This condition may of course be expressed in elementary terms, as the assertion that the internal category in $\mathcal{S}/C_0$ derived by the Grothendieck construction from the right action of $\mathbb{D}$ on $(T_0 \to C_0 \times D_0)$ is filtered.) Clearly, the class of such profunctors is closed under composition, so the right torsors are the 1-cells of a locally full sub-bicategory $\mathfrak{rTors}_{\mathcal{S}}$ of $\mathfrak{Prof}_{\mathcal{S}}$, which is dually equivalent to the full sub-2-category $\mathfrak{DTop}/\mathcal{S}$ of $\mathfrak{Top}/\mathcal{S}$ whose objects are the toposes $[\mathbb{C}, \mathcal{S}]$. And this equivalence identifies the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{DTop}/\mathcal{S}$ of 2.3.22 with the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{rTors}_{\mathcal{S}}^{\mathrm{op}}$ which sends an internal category $\mathbb{C}$ to itself and an internal functor f to $f^\bullet$ (cf. 2.7.4; it follows from the dual of 2.7.4(i) that profunctors of the form $f^\bullet$ are right torsors). We note that this functor is locally full and faithful, by 2.7.4(ii), and hence so is the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{DTop}/\mathcal{S}$; that is, for any two internal categories $\mathbb{C}$ and $\mathbb{D}$, the functor

$$\mathfrak{Cat}(\mathcal{S})\,(\mathbb{C}, \mathbb{D}) \longrightarrow \mathfrak{Top}/\mathcal{S}\,([\mathbb{C}, \mathcal{S}], [\mathbb{D}, \mathcal{S}])$$

is full and faithful.

We next discuss two important examples of the application of Diaconescu's Theorem, in which we shall presume familiarity with material from Chapter D5 (specifically, from Section D5.2 for the first example and from D5.4 for the second). The reason for placing these examples here is that they are both of fundamental importance in the theory of classifying toposes which we shall expound in Section B4.2.

Example 3.2.9 For the first one, we suppose our base topos $\mathcal{S}$ has a natural number object N: let C denote the generic finite cardinal in $\mathcal{S}/N$ (cf. A2.5.14) and let $\mathbb{S}_f$ be the internal full subcategory $\mathbb{S}[C]$ as defined in 2.3.5. Thus the object of objects of $\mathbb{S}_f$ is N, and its object of morphisms is (the domain of) the exponential $\pi_2^*(C)^{\pi_1^*(C)}$ in $\mathcal{S}/N \times N$. Since inverse image functors preserve natural number objects (A2.5.6(i)) and commute with exponentiation to the power of a finite cardinal (D5.2.11), we note that we have $p^*(\mathbb{S}_f) \cong \mathbb{E}_f$ for any $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$. Thus, once we have identified the $\mathbb{S}_f$-torsors in $\mathcal{S}$, for an arbitrary topos $\mathcal{S}$, we shall have identified the category $\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{S}_f, \mathcal{S}])$ for an arbitrary $\mathcal{S}$-topos $\mathcal{E}$. Now the indexed category corresponding to $\mathbb{S}_f$ is the full subcategory of $\mathbb{S}$ whose I-indexed families of objects are the finite cardinals in $\mathcal{S}/I$; it follows from D5.2.7 that this category is cocartesian, and so by 3.2.5 $\mathbb{S}_f$-torsors correspond to cartesian $\mathcal{S}$-indexed functors $\mathbb{S}_f^{\mathrm{op}} \to \mathbb{S}$.

Now $\mathbb{S}_f$ is the free category with finite $\mathcal{S}$-indexed coproducts generated by the terminal category $\mathbf{1}$ (cf. 1.4.17); but it is in fact the free cocartesian category generated by $\mathbf{1}$ – this can be shown using the facts that it already has coequalizers (D5.2.5) and its objects are (internally) projective (D5.2.9(ii)). Hence cartesian functors $\mathbb{S}_f^{\mathrm{op}} \to \mathbb{S}$ are equivalent to arbitrary functors $\mathbf{1} \to \mathbb{S}$, i.e. to objects of $\mathcal{S}$. We may also derive this equivalence more explicitly, as follows.

If A is any object of $\mathcal{S}$, then it is clear that the $\mathcal{S}$-indexed functor which sends a cardinal $[p]$ in $\mathcal{S}/I$ to $(I^*A)^{[p]}$ is cartesian. On the other hand, if $B \to N$ is the object of $\mathcal{S}/N$ underlying an $\mathbb{S}_f$-torsor, then preservation of finite products forces $o^*(B) \cong 1$ and $s^*(B) \cong A \times B$, where $A = (so)^*(B)$; so by the uniqueness theorem D5.1.9 we have $B \cong LA$, where LA is the list object over A (cf. A2.5.17). Hence every cartesian functor $\mathbb{S}_f^{\mathrm{op}} \to \mathbb{S}$ is of the form $[p] \mapsto (I^*A)^{[p]}$. Moreover, an $\mathcal{S}$-indexed natural transformation between two functors of this kind, corresponding to objects A and A', is necessarily induced by a (unique) morphism $A \to A'$ in $\mathcal{S}$, as may be seen by applying it to the element of $(A^*A)^{[so]}$ which corresponds to the diagonal map $A \rightarrowtail A \times A$. So we deduce that $\mathbf{Tors}(\mathbb{S}_f, \mathcal{S})$ is equivalent to $\mathcal{S}$ itself, and hence *for any $\mathcal{S}$-topos $\mathcal{E}$, we have*

$$\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{S}_f, \mathcal{S}]) \simeq \mathcal{E};$$

that is, $[\mathbb{S}_f, \mathcal{S}]$ is a representing object for the forgetful functor $\mathfrak{Top}/\mathcal{S}^{\mathrm{op}} \to \mathfrak{CAT}$ which sends a topos to its underlying category, and a geometric morphism to its inverse image functor.

Example 3.2.10 Let I be an object of $\mathcal{S}$, and let $K(I)$ denote the object of K-finite subobjects of I, as constructed in D5.4.1. Since $K(I)$ is a subobject of the power object PI, it carries a canonical partial ordering, i.e. we may regard it as the object of objects of an internal poset $\mathbb{K}(I)$. Once again, it follows from D5.4.12 that we have $p^*(\mathbb{K}(I)) \cong \mathbb{K}(p^*(I))$ for any geometric morphism $p: \mathcal{E} \to \mathcal{S}$, and so to identify the geometric morphisms over $\mathcal{S}$ into $[\mathbb{K}(I), \mathcal{S}]$ it suffices to identify the $\mathbb{K}(I)$-torsors in $\mathcal{S}$ itself. Also, $\mathbb{K}(I)$ is cocartesian (since it is a join-semilattice), so these torsors correspond to cartesian $\mathcal{S}$-indexed functors $\mathbb{K}(I)^{\mathrm{op}} \to \mathbb{S}$. Since $\mathbb{K}(I)$ is a poset, any such functor must take values in the subfibration $\mathbb{S}\mathsf{ub}$ of subterminal objects of $\mathbb{S}$ (cf. 1.2.2(g)); but since $\mathcal{S}$ is a topos, the latter is equivalent to $[\mathbf{\Omega}]$ where $\mathbf{\Omega}$ is the internal poset $(\Omega_1 \rightrightarrows \Omega)$ (cf. 2.3.8(a)). Thus, by 2.3.3, $\mathbb{K}(I)$-torsors correspond to internal cartesian functors (that is, meet-semilattice homomorphisms) $\mathbb{K}(I)^{\mathrm{op}} \to \mathbf{\Omega}$; but $\mathbb{K}(I)$ is the free join-semilattice generated by I (D5.4.9), and so these in turn correspond to arbitrary morphisms $I \to \Omega$, that is to subobjects of I. It is also straightforward to verify, by chasing through the above equivalences, that morphisms of $\mathbb{K}(I)$-torsors correspond to inclusions between subobjects of I. Thus we have shown that *for any object I of $\mathcal{S}$, we have an equivalence*

$$\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{K}(I), \mathcal{S}]) \simeq \mathrm{Sub}_{\mathcal{E}}(p^*(I))$$

which is natural in $\mathcal{E}$.

Remark 3.2.11 As a particular case of 3.2.10, we note that the functor category $[\mathbf{2}, \mathcal{S}]$ (where $\mathbf{2}$ denotes the finite category $(\bullet \to \bullet)$, identified with an internal poset in $\mathcal{S}$ by the method of 2.3.12(b)) is a classifying topos for subterminal objects in $\mathcal{S}$-toposes, since $K(1) \cong 1 \amalg 1$ by D5.4.4(i). This category is

often called the *Sierpiński topos* over $\mathcal{S}$: in the case $\mathcal{S} = \mathbf{Set}$, it may be identified with the topos of sheaves on the *Sierpiński space* S, i.e. the two-point space $\{0,1\}$ with $\{1\}$ (but not $\{0\}$) open.

In fact, in this case the equivalence $\mathfrak{Top}/\mathcal{S}\,(\mathcal{E},[\mathbf{2},\mathcal{S}]) \simeq \mathrm{Sub}_{\mathcal{E}}(1)$ may easily be verified without appealing to Diaconescu's Theorem. For $[\mathbf{2},\mathcal{S}]$ is the topos obtained by glueing along the identity functor $\mathcal{S} \to \mathcal{S}$ (cf. A2.1.12), and so by A4.5.6 it has complementary open and closed subtoposes, both isomorphic to $\mathcal{S}$. Hence a geometric morphism $\mathcal{E} \to [\mathbf{2},\mathcal{S}]$ induces, by pullback, a corresponding decomposition of $\mathcal{E}$ into complementary open and closed subtoposes. Conversely, given an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ equipped with a distinguished subterminal object U, we may identify $\mathcal{E}$ with a topos obtained by glueing as in A4.5.6, and then 'glue together' the morphisms $\mathbf{sh}_{o(U)}(\mathcal{E}) \to \mathcal{E} \to \mathcal{S}$ and $\mathbf{sh}_{c(U)}(\mathcal{E}) \to \mathcal{E} \to \mathcal{S}$ into a single morphism $\mathcal{E} \to [\mathbf{2},\mathcal{S}]$. We leave to the reader the verification that these two constructions are inverse to each other up to isomorphism.

An important application of Diaconescu's Theorem is the result that geometric morphisms of the form $[\mathbb{C},\mathcal{S}] \to \mathcal{S}$ have pullbacks in $\mathfrak{Top}$. At first sight, this appears to be a straightforward consequence of the theorem:

Corollary 3.2.12 *Let $\mathbb{C}$ be an internal category in a topos $\mathcal{S}$, and $p\colon \mathcal{E} \to \mathcal{S}$ a geometric morphism. Then the diagram*

$$\begin{tikzcd} {[p^*(\mathbb{C}),\mathcal{E}]} \arrow[r, "p^{\mathbb{C}}"] \arrow[d] & {[\mathbb{C},\mathcal{S}]} \arrow[d] \\ \mathcal{E} \arrow[r, "p"] & \mathcal{S} \end{tikzcd}$$

constructed in the proof of 3.2.7 *is a product diagram in* $\mathfrak{Top}/\mathcal{S}$.

Proof For any $\mathcal{S}$-topos $q\colon \mathcal{F} \to \mathcal{S}$, pairs of geometric morphisms $f\colon \mathcal{F} \to \mathcal{E}$, $g\colon \mathcal{F} \to [\mathbb{C},\mathcal{S}]$ over $\mathcal{S}$ correspond, up to equivalence of categories, to pairs (f,G) where G is a torsor over $\mathbb{C}$ (that is, over $q^*(\mathbb{C}) \cong f^*p^*(\mathbb{C})$) in $\mathcal{F}$. But these in turn correspond to pairs (f,h), where h is a geometric morphism $\mathcal{F} \to [p^*(\mathbb{C}),\mathcal{E}]$ over $\mathcal{E}$. □

As a special case of 3.2.12, we note

Corollary 3.2.13 *The functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$ of* 2.3.22 *preserves finite products.*

Proof It clearly preserves the terminal object; to prove that it preserves binary products, take $\mathcal{E}$ to be a topos of the form $[\mathbb{D},\mathcal{S}]$ in 3.2.12. Then $[p^*(\mathbb{C}),\mathcal{E}]$ is equivalent to $[\mathbb{C} \times \mathbb{D},\mathcal{S}]$, since the objects of both categories may be identified with discrete opfibrations over $\mathbb{C} \times \mathbb{D}$ by 2.5.2 and 2.5.3. □

However, if we wish to prove that the square in the statement of 3.2.12 is a (2-categorical) pullback in $\mathfrak{Top}$, and not merely in $\mathfrak{Top}/\mathcal{S}$, we have to consider non-invertible geometric transformations between geometric morphisms into $\mathcal{S}$, and such 2-cells do not live in the 2-category $\mathfrak{Top}/\mathcal{S}$. To get round this difficulty, we need to borrow a result from Section B3.4 (see 3.4.2 below): namely, that the 2-category $\mathfrak{Top}$ admits tensors with $\mathbf{2}$. That is, for any topos $\mathcal{E}$ there is a topos $\mathbf{2}\otimes\mathcal{E}$ (in fact it is the Sierpiński topos $[\mathbf{2},\mathcal{E}]$, cf. 3.2.11) and a natural equivalence

$$\mathfrak{Top}\,(\mathbf{2}\otimes\mathcal{E},\mathcal{F}) \simeq [\mathbf{2},\mathfrak{Top}\,(\mathcal{E},\mathcal{F})]$$

for all toposes $\mathcal{F}$.

Corollary 3.2.14 *For any geometric morphism $p\colon\mathcal{E}\to\mathcal{S}$ and any internal category $\mathbb{C}$ in $\mathcal{S}$, the square in the statement of* 3.2.12 *is a pullback in* $\mathfrak{Top}$.

Proof Given a topos $\mathcal{F}$ and a pair of geometric morphisms $f\colon\mathcal{F}\to\mathcal{E}$, $g\colon\mathcal{F}\to[\mathbb{C},\mathcal{S}]$ satisfying $pf \cong \ulcorner\mathbb{C}\urcorner g$, we obtain a geometric morphism $h\colon\mathcal{F}\to[p^*(\mathbb{C}),\mathcal{E}]$ as in 3.2.12. Given two such pairs (f_0,g_0) and (f_1,g_1) together with geometric transformations $\alpha\colon f_0\to f_1$ and $\beta\colon g_0\to g_1$ satisfying $p\circ\alpha = \ulcorner\mathbb{C}\urcorner\circ\beta$, we may regard the data as a compatible pair of morphisms from $\mathbf{2}\otimes\mathcal{F}$ to the diagram, and hence obtain a morphism $\mathbf{2}\otimes\mathcal{F}\to[p^*(\mathbb{C}),\mathcal{E}]$, or equivalently a geometric transformation between the two morphisms $h_0,h_1\colon\mathcal{F}\rightrightarrows[p^*(\mathbb{C}),\mathcal{E}]$. It is now straightforward to verify that these constructions yield a functor

$$\mathfrak{Top}\,(\mathcal{F},\mathcal{E})\times_{\mathfrak{Top}\,(\mathcal{F},\mathcal{S})}\mathfrak{Top}\,(\mathcal{F},[\mathbb{C},\mathcal{S}]) \longrightarrow \mathfrak{Top}\,(\mathcal{F},[p^*(\mathbb{C}),\mathcal{E}])$$

which is inverse up to natural isomorphism to the functor obtained by composing with the left and top edges of the square of 3.2.12. □

Suggestion for further reading: Diaconescu [280].

B3.3 Giraud's Theorem

In this section we shall prove the fundamental representation theorem for bounded $\mathcal{S}$-toposes (again due to R. Diaconescu [280]), which says that a geometric morphism $p\colon\mathcal{E}\to\mathcal{S}$ is bounded iff it can be factored as a composite

$$\mathcal{E} \xrightarrow{\;i\;} [\mathbb{C},\mathcal{S}] \longrightarrow \mathcal{S}$$

where $\mathbb{C}$ is an internal category in $\mathcal{S}$ and i is a geometric inclusion. Bearing in mind the identification of inclusions $\mathcal{E}\to\mathcal{F}$ with local operators on $\mathcal{F}$, established in Section A4.4, this is tantamount to saying that any bounded $\mathcal{S}$-topos $\mathcal{E}$ is equivalent to the topos of '$\mathcal{S}$-valued sheaves' on an 'internal site' in $\mathcal{S}$. The representation theorem is commonly known as Giraud's Theorem, because J. Giraud [405] proved a representation theorem of this kind for the case $\mathcal{S} = \mathbf{Set}$;

but it is not really the $\mathcal{S}$-indexed analogue of Giraud's original theorem, because Giraud did not have the elementary concept of topos available, and therefore sought a characterization of categories of sheaves on sites amongst arbitrary categories, rather than amongst toposes defined over **Set**. (In modern terminology, his theorem is essentially the equivalence of (ii) and (vii) in Theorem C2.2.8 – whereas the theorem we are about to prove is (the analogue over an arbitrary base of) the equivalence of (ii) and (v).) We shall return to this topic at the end of the section, and show how a theorem more like Giraud's original may be extracted from the one we have proved here.

For now, let $p\colon \mathcal{E} \to \mathcal{S}$ be a bounded geometric morphism, and let $(g\colon G \to p^*I) \in \mathrm{ob}\,\mathcal{E}^I$ (for some $I \in \mathrm{ob}\,\mathcal{S}$) be a separating family for the $\mathcal{S}$-indexed category $\mathbb{E}$. We take our internal category $\mathbb{C}$ to be the opposite of the internal full subcategory $\mathbb{E}[g]$, defined as in 2.3.5: explicitly, $C_0 = I$, and the object of morphisms and domain and codomain maps of $\mathbb{C}$ are given by the pullback

$$\begin{tikzcd} C_1 \arrow[r] \arrow[d, "{(d_0,d_1)}"] & p_*E \arrow[d, "{(p_*a,p_*b)}"] \\ C_0 \times C_0 \arrow[r, "\eta"] & p_*p^*I \times p_*p^*I \end{tikzcd}$$

where $(a,b)\colon E \to p^*I \times p^*I$ is the exponential $\pi_2^*(g)^{\pi_1^*(g)}$ in $\mathcal{E}/p^*I \times p^*I$, and η is the unit of $(p^* \dashv p_*)$. Then, by 2.3.13, the $\mathcal{S}$-indexed full embedding $[\mathbb{C}]^{\mathrm{op}} \to \mathbb{E}$ corresponds to an internal diagram F of shape $p^*(\mathbb{C})^{\mathrm{op}}$ in $\mathcal{E}$: explicitly, $(F_0 \to C_0)$ is $g\colon G \to p^*I$, and the right action of $p^*\mathbb{C}$ on F_0 is given as follows. We have a commutative diagram

$$\begin{tikzcd} F_0 \times_{p^*C_0} p^*C_1 \arrow[r] \arrow[d] & p^*C_1 \arrow[r] \arrow[d, "d_0"] & p^*p_*E \arrow[r, "\epsilon"] \arrow[d, "p^*p_*a"] & E \arrow[d, "a"] \\ G \arrow[r, "g"] & p^*I \arrow[r, "p^*\eta"] & p^*p_*p^*I \arrow[r, "\epsilon"] & p^*I \end{tikzcd}$$

(where ϵ is the counit of $(p^* \dashv p_*)$), and hence a morphism $F_0 \times_{p^*C_0} p^*C_1 \to G \times_{p^*I} E$. But the codomain of this morphism is (the domain of) the product $\pi_1^*(g) \times \pi_2^*(g)^{\pi_1^*(g)}$ in $\mathcal{E}/p^*I \times p^*I$, and so we have an evaluation map from here to $\pi_2^*(g)$, which we compose with the projection $\Sigma_{p^*I \times p^*I}(\pi_2^*(g)) \cong p^*I \times G \to G$.

Lemma 3.3.1 *The diagram F just defined is a $\mathbb{C}$-torsor in $\mathcal{E}$.*

Proof Applying 2.4.3 to the terminal object of $\mathcal{E}^1$, we obtain a morphism $x\colon J \to I$ in $\mathcal{S}$ and an epimorphism $\Sigma_J x^*(g) \twoheadrightarrow 1$ in $\mathcal{E}$; but the latter factors

through $G \to 1$, and so it too must be epic. Thus the first condition of 2.6.2(a) is verified for $\mathbb{F}$.

Next, we apply 2.4.3 to the object $(g \times g \colon G \times G \to p^*I \times p^*I)$ of $\mathcal{E}^{I \times I}$. We obtain a commutative diagram of the form

$$\begin{tikzcd} G \arrow[d, "g"] & H \arrow[l, "q_3"'] \arrow[r, two heads, "(q_1, q_2)"] \arrow[d, "h"] & G \times G \arrow[d, "g \times g"] \\ p^*I & p^*K \arrow[l, "p^*y_3"'] \arrow[r, "(p^*y_1, p^*y_2)"] & p^*I \times p^*I \end{tikzcd}$$

where the left-hand square is a pullback. We shall show that the epimorphism (q_1, q_2) factors through the morphism $P_{\mathbb{F}} \to F_0 \times F_0 = G \times G$ of 2.6.2(a), so that the latter is also epic.

Now q_1 induces a morphism $y_3^*(g) \to y_1^*(g)$ in $\mathcal{E}^K$, and hence a morphism $z \colon K \to C_1$ in $\mathcal{S}$ satisfying $d_0 z = y_3$ and $d_1 z = y_1$. And this in turn induces a morphism

$$r_1 = (q_3, p^*(z)h) \colon H \longrightarrow F_0 \times_{p^*C_0} p^*C_1 = F_1$$

in $\mathcal{E}$ satisfying $d_0 r_1 = q_3$ and $d_1 r_1 = q_1$. Similarly, we obtain $r_2 \colon H \to F_1$ satisfying $d_0 r_2 = q_3$ and $d_1 r_2 = q_2$; so the pair (r_1, r_2) is the required factorization of (q_1, q_2) through $P_{\mathbb{F}}$.

For the third condition of 2.6.2(a), we similarly start by applying 2.4.3 to the object $(R_{\mathbb{F}} \to p^*(R_{\mathbb{C}}))$ of $\mathcal{E}^{R_{\mathbb{C}}}$, where (as in 2.6.1(c)) R denotes the 'object of parallel pairs' functor. We obtain a diagram

$$\begin{tikzcd} G \arrow[d, "g"] & B \arrow[l, "s_3"'] \arrow[r, two heads, "t"] \arrow[d, "b"] & R_{\mathbb{F}} \arrow[d] \\ p^*I & p^*L \arrow[l, "w_3"'] \arrow[r, "u"] & p^*R_{\mathbb{C}} \end{tikzcd}$$

where, once again, the left-hand square is a pullback. We shall write w_1 and w_2 for the composites of u with the 'domain' and 'codomain' morphisms $R_{\mathbb{C}} \rightrightarrows C_0 = I$, and similarly s_1 and s_2 for the composites of t with $R_{\mathbb{F}} \rightrightarrows F_0 = G$. Now u corresponds to a parallel pair of morphisms $w_2^*(g) \rightrightarrows w_1^*(g)$ in $\mathcal{E}^L$, and t corresponds to a morphism from $w_3^*(g)$ to the equalizer of this pair. Thus we have a diagram

$$w_3^*(g) \longrightarrow w_2^*(g) \rightrightarrows w_1^*(g)$$

of shape $\mathcal{T}^{\mathrm{op}}$ (where $\mathcal{T}$ is the finite category displayed in 2.6.1(e)) in $\mathcal{E}^L$, which corresponds to a morphism $v\colon L \to T_{\mathbb{C}}$ such that

$$\begin{tikzcd} L \arrow[r, "v"] \arrow[dr, "u"'] & T_{\mathbb{C}} \arrow[d] \\ & R_{\mathbb{C}} \end{tikzcd}$$

commutes. Now the pair $(s_3, p^*(v)b)\colon B \to F_0 \times_{p^*C_0} p^*T_{\mathbb{C}} \cong T_{\mathbb{F}}$ similarly yields a factorization of t through $T_{\mathbb{F}} \to R_{\mathbb{F}}$; so the latter must be epic. □

Remark 3.3.2 The proof of 3.3.1 may be somewhat simplified if we assume that our separating family is 'closed under subobjects' in the sense that, given any $x\colon J \to I$ and any subobject $S \rightarrowtail x^*G$, there exists a morphism $y\colon J \to I$ such that

$$\begin{tikzcd} S \arrow[r] \arrow[d] & G \arrow[d, "g"] \\ p^*J \arrow[r, "p^*y"] & p^*I \end{tikzcd}$$

is a pullback. We note that this condition is always satisfied by the separating family constructed from a bound for $\mathcal{E}$ over $\mathcal{S}$ as in the proof of 3.1.6(ii) $\Rightarrow$ (i). (For an account of the proof in this particular case, see [504].)

Still further simplification is possible when the base topos $\mathcal{S}$ has a natural number object: for in this case we may assume that our separating family is closed under finite products as well as subobjects, by the device of replacing a separating family $(G \to p^*I)$ closed under subobjects by the family $(LG \to L(p^*I) \cong p^*(LI))$ where LG is the list object over G as constructed in A2.5.17. Then we see that our internal full subcategory $\mathbb{C}^{\mathrm{op}}$ has $\mathcal{S}$-indexed finite limits and the inclusion functor $\mathbb{C}^{\mathrm{op}} \to \mathbf{E}$ is cartesian; so the result of 3.3.1 follows from 3.2.5. The same argument may be employed, even if $\mathcal{S}$ does not have a natural number object, in the case when p is localic (cf. 3.3.5(iii) below).

However, for the next step in the argument there seems to be no substitute for hard calculation.

Lemma 3.3.3 *The geometric morphism $i\colon \mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ which corresponds to the torsor F of* 3.3.1 *is an inclusion.*

Proof We first observe that the direct image functor i_* may be described as follows: given an object A of $\mathcal{E}$, we form the exponential $(I^*A)^g$ in $\mathcal{E}^I$, apply p_* to it and pull it back along the unit $I \to p_*p^*I$ to obtain an object $(B_0 \to I)$ of $\mathcal{S}/I$. It is easy to see that this object admits a left action of $\mathbb{C}$ induced by the

composition map $\pi_2^* g^{\pi_1^* g} \times ((I \times I)^* A)^{\pi_2^* g} \to ((I \times I)^* A)^{\pi_1^* g}$ in $\mathcal{E}^{I \times I}$, yielding a diagram B of shape $\mathbb{C}$ which we take to be $i_*(A)$. For, given an arbitrary diagram D of shape $\mathbb{C}$, morphisms $D_0 \to B_0$ over I correspond to morphisms $G \times_{p^* I} p^* D_0 \to A$ in $\mathcal{E}$, and it is straightforward to verify that a morphism $D_0 \to B_0$ is equivariant for the left $\mathbb{C}$-actions iff the corresponding morphism $G \times_{p^* I} p^* D_0 \to A$ factors through the coequalizer $G \times_{p^* I} p^* D_0 \twoheadrightarrow F \otimes_{p^* \mathbb{C}} p^* D = i^* D$.

We now show that i_* is faithful. Suppose given a parallel pair $f, f' \colon A \rightrightarrows A'$ in $\mathcal{E}$ with $f \neq f'$. Then we can find $x \colon J \to I$ in $\mathcal{S}$ and a morphism $h \colon x^* g \to J^* A$ in $\mathcal{E}^J$ with $(J^* f)h \neq (J^* f')h$. But morphisms $x^* g \to J^* A$ in $\mathcal{E}^J$ correspond to morphisms $J \to B_0$ over I (where B_0 is defined as above), and so we have such a morphism whose composites with $i_* f$ and $i_* f'$ are different; thus $i_* f \neq i_* f'$.

To show that i_* is full as well as faithful, we observe that a morphism $f \colon A \to A'$ in $[\mathbb{C}, \mathcal{S}]$ induces, for every $x \colon J \to I$ and every $u \colon \Sigma_J x^*(g) \to A$, a morphism $\overline{f}(u) \colon \Sigma_J x^*(g) \to A'$, in a manner which is 'natural' with respect to morphisms between pullbacks of g. Now we may use 2.4.3 to construct a diagram

$$\begin{tikzcd}
M \arrow[r, two heads, "v"] \arrow[d, "m"] & R \arrow[r, shift left, "a"] \arrow[r, shift right, "b"'] \arrow[d] & H \arrow[r, two heads, "u"] \arrow[d, "h"] & A \\
p^* K \arrow[r, "p^* z"] \arrow[d, "p^* y"] & p^* J \times p^* J \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] & p^* J \arrow[d, "p^* x"] & \\
p^* I & & p^* I &
\end{tikzcd}$$

where (a, b) is the kernel-pair of u, and h and m are the pullbacks of g along $p^* x$ and $p^* y$ respectively. The argument above yields a morphism $\overline{f}(u) \colon H \to A'$, and the naturality condition tells us that this morphism has equal composites with av and bv, so it factors through their coequalizer, which is u. Thus we obtain a morphism $\check{f} \colon A \to A'$; and a further application of the same technique shows that, for any $u' \colon \Sigma_{J'} x'^*(g) \to A$, $\overline{f}(u')$ is the composite $\check{f} u'$, and hence $i_*(\check{f}) = f$. □

Combining 3.3.1 and 3.3.3, we have proved

Theorem 3.3.4 (Giraud's Theorem) *For a geometric morphism $p \colon \mathcal{E} \to \mathcal{S}$, the following are equivalent:*

(i) *p is bounded.*

(ii) *There exists an internal category $\mathbb{C}$ in $\mathcal{S}$ and an inclusion $\mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ in $\mathfrak{Top}/\mathcal{S}$.*

Proof (i) $\Rightarrow$ (ii) follows directly from 3.3.1 and 3.3.3. For the converse, we use the fact that $[\mathbb{C}, \mathcal{S}]$ is bounded over $\mathcal{S}$ by 3.2.1, together with 2.4.2(d). $\square$

It is of course possible to introduce a notion of 'internal coverage' on an internal category $\mathbb{C}$, corresponding to the external notion introduced in A2.1.9, and to say what it means for a diagram of shape $\mathbb{C}^{\mathrm{op}}$ to be a sheaf for such a coverage. Then one may prove that every subtopos of a diagram topos $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ is the category of T-sheaves for a suitable coverage T, and thus deduce from 3.3.4 that every bounded $\mathcal{S}$-topos is of this form. However, we shall find it convenient to postpone consideration of the notion of internal coverage until Section C2.4.

As a particular case of Giraud's Theorem, we note

Corollary 3.3.5 *For a geometric morphism $p: \mathcal{E} \to \mathcal{S}$, the following are equivalent:*

(i) *p is localic.*

(ii) *There exists an internal preorder $\mathbb{P}$ in $\mathcal{S}$ and an inclusion $\mathcal{E} \to [\mathbb{P}, \mathcal{S}]$ of $\mathcal{S}$-toposes.*

(iii) *There exists an internal complete Heyting algebra $\mathbb{H}$ in $\mathcal{S}$ and an inclusion $\mathcal{E} \to [\mathbb{H}^{\mathrm{op}}, \mathcal{S}]$ of $\mathcal{S}$-toposes.*

Proof (ii) $\Rightarrow$ (i) follows from 3.2.2 and the facts, noted in A4.6.2(a) and (e), that inclusions are localic and composites of localic morphisms are localic.

(iii) $\Rightarrow$ (ii) is trivial.

(i) $\Rightarrow$ (iii): If 1 is a bound for $\mathcal{E}$ over $\mathcal{S}$, then by 3.1.6 we obtain a separating family for $\mathbb{E}$ by taking the left vertical map in the pullback

$$
\begin{tikzcd}
H \arrow[r, tail] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\
p^* p_*(\Omega) \arrow[r, "\epsilon_\Omega"] & \Omega
\end{tikzcd}
$$

considered as an object of $\mathcal{E}^{p_*(\Omega)}$. It is then easy to see that the internal category $\mathbb{C}$ constructed from this separating family as above is simply $p_*(\boldsymbol{\Omega})^{\mathrm{op}}$, where $\boldsymbol{\Omega}$ is the internal poset constructed from the subobject classifier of $\mathcal{E}$ as in A1.6.3. But the latter is an internal Heyting algebra by A1.6.3, and $\mathcal{E}$-complete by 2.3.8(a); and both these properties are preserved by the functor p_*. $\square$

We note that the particular case of 3.3.5 where $\mathcal{S}$ = **Set** appears, with a different proof, as C1.4.7.

One of the most important applications of Giraud's Theorem is to the construction of pullbacks in $\mathfrak{Top}$:

Proposition 3.3.6 *Let $p: \mathcal{E} \to \mathcal{S}$ and $q: \mathcal{F} \to \mathcal{S}$ be geometric morphisms with common codomain, such that p is bounded (resp. localic). Then there exists a*

pullback square

$$\begin{tikzcd} \mathcal{G} \arrow[r, "f"] \arrow[d, "g"] & \mathcal{E} \arrow[d, "p"] \\ \mathcal{F} \arrow[r, "q"] & \mathcal{S} \end{tikzcd}$$

in $\mathfrak{Top}$, *in which* g *is bounded (resp. localic).*

Proof By 3.3.4 (resp. 3.3.5) we may factor p as

$$\mathcal{E} \xrightarrow{i} [\mathbb{C}, \mathcal{S}] \longrightarrow \mathcal{S},$$

where $\mathbb{C}$ is an internal category (resp. an internal preorder) in $\mathcal{S}$ and i is an inclusion. By 3.2.14 the pullback of $[\mathbb{C}, \mathcal{S}]$ along q exists and is of the form $[q^*(\mathbb{C}), \mathcal{S}]$; and clearly $q^*(\mathbb{C})$ is a preorder if $\mathbb{C}$ is. So it suffices to show that the inclusion i can be pulled back along $q^{\mathbb{C}}$, and that its pullback is an inclusion; but this was done in A4.5.14(e). □

As at the end of Section B3.1, we should emphasize that although the boundedness restriction in 3.3.6 is a real one, it is not in practice very irksome, since we know so few examples of unbounded geometric morphisms. Indeed, we do not know any example of a pair of unbounded morphisms whose pullback is definitely known not to exist in $\mathfrak{Top}$.

For future reference, we note

Lemma 3.3.7 *Let* $q: \mathcal{F} \to \mathcal{S}$ *be a hyperconnected geometric morphism. Then the pullback of* q *along any bounded geometric morphism* $p: \mathcal{E} \to \mathcal{S}$ *is hyperconnected.*

Proof As before, it suffices to consider separately the cases when p is of the form $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$, and when p is an inclusion. In the former case, since the inverse image of the pullback $q^{\mathbb{C}}$ of q is simply q^* applied to diagrams of shape $\mathbb{C}$ in $\mathcal{S}$, it is easy to see that it is full and faithful and that its image is closed under subobjects. The latter case was dealt with in A4.6.11. □

In particular, it follows from 3.3.6 and 3.3.7 that the hyperconnected–localic factorization of an arbitrary geometric morphism, as constructed in A4.6.5, is stable under pullback along arbitrary bounded morphisms. In C2.4.11 we give an entirely different proof that bounded hyperconnected morphisms are stable under pullback along arbitrary geometric morphisms.

Since any bounded geometric morphism $p: \mathcal{E} \to \mathcal{S}$ can be pulled back along itself, we can form the diagonal map $\Delta_p: \mathcal{E} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$. In Part C we shall often have occasion to consider such diagonal morphisms, and the following result will be of use.

Proposition 3.3.8

(i) *For any bounded geometric morphism p, the diagonal Δ_p is localic.*

(ii) *If p is localic, then Δ_p is an inclusion.*

(iii) *If p is an inclusion, then Δ_p is an equivalence.*

Proof It is convenient to prove the third assertion first. But (iii) is immediate from the construction of pullbacks of inclusions given in A4.5.14(e); for the pullback of any inclusion along itself is clearly an equivalence. (Alternatively, we could appeal to the fact that inclusions are fully monic in $\mathfrak{Top}$, in the sense of 1.1.19.)

(i) We may factor p, as usual, as $\mathcal{E} \to [\mathbb{C}, \mathcal{S}] \to \mathcal{S}$, where the first factor is an inclusion. For the second factor, we know by 3.2.13 that $[\mathbb{C}, \mathcal{S}] \times_{\mathcal{S}} [\mathbb{C}, \mathcal{S}] \simeq [\mathbb{C} \times \mathbb{C}, \mathcal{S}]$, and it is clear that the diagonal map of this factor is induced by the diagonal functor $\mathbb{C} \to \mathbb{C} \times \mathbb{C}$. But the latter functor is faithful, so by (the internal version of) A4.6.2(c) it induces a localic geometric morphism. Now we have a pullback square

$$\begin{tikzcd} \mathcal{E} \times_{[\mathbb{C},\mathcal{S}]} \mathcal{E} \arrow[r] \arrow[d] & {[\mathbb{C},\mathcal{S}]} \arrow[d] \\ \mathcal{E} \times_{\mathcal{S}} \mathcal{E} \arrow[r] & {[\mathbb{C},\mathcal{S}] \times_{\mathcal{S}} [\mathbb{C},\mathcal{S}]} \end{tikzcd}$$

where the right vertical arrow is the diagonal of $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$, so the left vertical arrow is also localic by 3.3.6. But the diagonal of p is the composite of this morphism with the diagonal of $\mathcal{E} \to [\mathbb{C}, \mathcal{S}]$, and we have already seen that the latter is an equivalence.

(ii) Similarly, if p is localic, we may factor it as $\mathcal{E} \to [\mathbb{P}, \mathcal{S}] \to \mathcal{S}$ where $\mathbb{P}$ is an internal preorder in $\mathcal{S}$, by 3.3.5. In this case, the diagonal functor $\mathbb{P} \to \mathbb{P} \times \mathbb{P}$ is full as well as faithful, so it follows from the internal version of A4.2.12(b) that the diagonal of $[\mathbb{P}, \mathcal{S}] \to \mathcal{S}$ is an inclusion. The rest of the proof is similar to that of (i). □

We shall eventually see (in C2.4.14) that the converse of 3.3.8(ii) is true. The converse of 3.3.8(iii) is, however, false: a counterexample is given by the morphism $\mathbf{Set}/X \to \mathbf{Sh}(X)$ of A4.2.4(e), for any non-discrete space X. When $\mathcal{S}$ has a natural number object, alternative proofs of both parts (i) and (ii) of 3.3.8 may be given using the notion of classifying topos developed in Section B4.2 below (and also in chapter D3). If $p\colon \mathcal{E} \to \mathcal{S}$ classifies a geometric theory $\mathbb{T}$ over $\mathcal{S}$, then $\mathcal{E} \times_{\mathcal{S}} \mathcal{E}$ classifies the theory of pairs of $\mathbb{T}$-models, and as a topos over it the diagonal may be viewed as classifying the theory of isomorphisms between the generic pair of $\mathbb{T}$-models – which is a propositional theory, and hence classified by a localic topos (4.2.12). Similarly, if p itself is localic, then we may take $\mathbb{T}$ to be a propositional theory, in which case freely adjoining an isomorphism between

two $\mathbb{T}$-models is equivalent to forcing the models to be the same – which, as we shall see in Section B4.2, can be done by an inclusion.

One might be tempted to conjecture after 3.3.8 that if p is (bounded and) hyperconnected than its diagonal should be a surjection. But this is false: consider the topos $[\mathcal{C}, \mathbf{Set}]$ where $\mathcal{C}$ is the free category generated by the directed graph

$$\bullet \rightleftarrows \bullet \; .$$

Then it follows easily from A4.6.9 that $[\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}$ is hyperconnected. However, neither of the two objects of $\mathcal{C}$ is a retract of the other, so it follows from A4.2.7(*b*) that the diagonal morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{C} \times \mathcal{C}, \mathbf{Set}]$ is not surjective.

We now return to the problem, mentioned at the beginning of this section, of formulating an $\mathcal{S}$-indexed version of Giraud's original theorem. In order to do this, we shall need to introduce the $\mathcal{S}$-indexed version of the notion of ∞-pretopos, which we met originally in Section A1.4.

Definition 3.3.9 Let $\mathcal{S}$ be a topos, and $\mathbb{E}$ an $\mathcal{S}$-indexed category. We say $\mathbb{E}$ is an *$\mathcal{S}$-indexed ∞-pretopos* if it satisfies the following conditions.

(i) Each fibre $\mathcal{E}^I$ is a pretopos, and the transition functors $x^* : \mathcal{E}^I \to \mathcal{E}^J$ are coherent functors.

(ii) Each $\mathcal{E}^I$ has coequalizers (of arbitrary pairs) which are stable under pullback and preserved by the functors x^*. Moreover, if

$$\begin{tikzcd} A' \arrow[r, shift left, "f'"] \arrow[r, shift right, "g'"'] \arrow[d, tail] & B' \arrow[d, tail] & \\ A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, two heads, "h"] & C \end{tikzcd}$$

is a commutative diagram in some $\mathcal{E}^I$ where the bottom row is a coequalizer and both the squares are pullbacks, then $B' \rightarrowtail B$ is the pullback along h of the image of the composite $B' \rightarrowtail B \twoheadrightarrow C$.

(iii) $\mathbb{E}$ is well-powered as an $\mathcal{S}$-indexed category.

(iv) $\mathbb{E}$ has $\mathcal{S}$-indexed coproducts which are disjoint and stable, in the sense defined in Section B1.4.

In the case $\mathcal{S} = \mathbf{Set}$, it is not hard to verify that the naive indexing of an ordinary category $\mathcal{E}$ is a $\mathbf{Set}$-indexed ∞-pretopos iff $\mathcal{E}$ is an ∞-pretopos as defined in Section A1.4. The only surprising element of the definition is the condition on arbitrary coequalizers in (ii), which was not present in our definition of an ∞-pretopos (though it follows from it, by A1.4.19; note that if the

pair (f, g) in the second condition of (ii) is an equivalence relation, then we have $B' \cong \exists_f g^*(B')$, from which $B' \cong h^*\exists_h(B')$ follows by Beck–Chevalley). Condition (ii) was also not present in Giraud's original theorem (he required merely that coequalizers of equivalence relations should exist and be stable under pullback – which is true in any pretopos). The reason why we have had to include it here is that we are not assuming that our base topos $\mathcal{S}$ has a natural number object, and so we cannot hope to construct the equivalence relation generated by an arbitrary parallel pair as we did in A1.4.19. (However, it would be sufficient to restrict the condition to coequalizers of reflexive pairs, since we already have well-behaved finite coproducts given by (i).) We note also that if $p\colon \mathcal{E} \to \mathcal{S}$ is a geometric morphism, then the indexing of $\mathcal{E}$ over $\mathcal{S}$ induced by p^* satisfies the conditions of 3.3.9: to verify that the second part of (ii) holds in any topos, note that the hypotheses imply that the classifying map $\chi\colon B \to \Omega$ of $B' \rightarrowtail B$ factors through h. Conversely, we have:

Lemma 3.3.10 *Let $\mathbb{E}$ be an $\mathcal{S}$-indexed ∞-pretopos with a separating family. Then $\mathbb{E}$ is an $\mathcal{S}$-indexed topos as defined in Section* B3.1.

Proof We note first that, by 1.4.12, $\mathbb{E}$ is equivalent to the $\mathcal{S}$-indexed category $(I \mapsto \mathcal{E}^1/p^*I)$ for some cartesian functor $p^*\colon \mathcal{S} \to \mathcal{E}^1$; moreover, we shall see shortly that $\mathbb{E}$ is locally small, so that p^* has a right adjoint by 2.2.5. Thus it is sufficient to show that the fibre $\mathcal{E}^1$ is a topos. However, our method for doing this will easily yield the extra information that the other fibres are also toposes, and the transition functors between them are logical; so we do not in fact need to appeal to 1.4.12.

To show that $\mathbb{E}$ is locally small, we use the $\mathcal{S}$-indexed version of the argument of A1.4.17: given objects A and B of fibres $\mathcal{E}^I$ and $\mathcal{E}^J$ respectively, the object of $\mathcal{S}/I \times J$ indexing morphisms $A \to B$ is a subobject of the object indexing subobjects of $\pi_1^*A \times \pi_2^*B$, corresponding to those subobjects $(a, b)\colon S \rightarrowtail \pi_1^*A \times \pi_2^*B$ which are graphs of morphisms, i.e. such that a is an isomorphism. To construct this subobject, we need to know that invertibility is definable in $\mathbb{E}$; and we cannot use 1.3.15 for this purpose, since we do not yet know that $\mathbb{E}$ is locally small. However, it is easy to see that in a well-powered indexed category the property of being invertible is definable for monomorphisms, since $x^*(m\colon A' \rightarrowtail A)$ is invertible iff x factors through the equalizer of the morphisms in $\mathcal{S}$ indexing m and 1_A; and an arbitrary morphism $f\colon A \to B$ in $\mathcal{E}^I$ is monic iff the diagonal $A \rightarrowtail A \times_B A$ is an isomorphism, so we may first construct the subobject $I' \rightarrowtail I$ which forces this to be true, and then construct $I'' \rightarrowtail I'$ forcing f itself to be an isomorphism.

$\mathbb{E}$ is also well-copowered, by an argument which we have seen before (for toposes) in 1.3.14: quotients of an object A of $\mathcal{E}^I$ correspond to subobjects of $A \times A$ which are equivalence relations, and we can construct an object of $\mathcal{S}/I$ which indexes the latter. Since it is also $\mathcal{S}$-cocomplete, we can therefore use the $\mathcal{S}$-indexed adjoint functor theorem to construct right adjoints for indexed functors defined on $\mathbb{E}$.

Thus, to show that $\mathcal{E}^1$ is cartesian closed, it suffices to show that for each object A of $\mathcal{E}^1$ the indexed functor which sends $B \in \text{ob } \mathcal{E}^I$ to $B \times I^*A$ is $\mathcal{S}$-cocontinuous. But conditions (i) and (ii) of 3.3.9 ensure that this functor preserves (fibrewise) finite colimits, so we need only check that it preserves $\mathcal{S}$-indexed coproducts. For this, we note that, given $x\colon J \to I$ and objects B, C of $\mathcal{E}^J, \mathcal{E}^I$ respectively, we have a pullback square

$$\begin{tikzcd} B \times x^*C \arrow[r, "\pi_1"] \arrow[d] & B \arrow[d, "\eta_B"] \\ x^*(\Sigma_x B \times C) \arrow[r] & x^*\Sigma_x B \end{tikzcd}$$

since x^* preserves finite products; hence by 1.4.10 the canonical morphism $\Sigma_x(B \times x^*C) \to \Sigma_x B \times C$ is an isomorphism. Specializing to the case when $C = I^*A$, we deduce that the indexed functor $(-) \times A$ preserves $\mathcal{S}$-indexed coproducts.

Thus $\mathcal{E}^1$ is cartesian closed; and repeating the argument in the context of $\mathcal{S}/I$-indexed categories will show that $\mathcal{E}^I$ is cartesian closed for every I. Moreover, we have verified the Frobenius reciprocity condition for the transition functors x^* and their left adjoints, so by A1.5.8 the x^* are cartesian closed functors.

To verify that the $\mathcal{E}^I$ have subobject classifiers, we consider the indexed functor $\mathbb{S}\mathrm{ub}_{\mathbb{E}}\colon \mathbb{E}^{\mathrm{op}} \to \mathbb{S}$ which sends an object A of $\mathcal{E}^I$ to the object of $\mathcal{S}/I$ indexing subobjects of A. We shall show that this functor is continuous, and so has an indexed left adjoint L; it is then straightforward to verify that $L^I(1_I)$ is a subobject classifier for $\mathcal{E}^I$, and since the L^I form an indexed functor it follows that the transition functors of $\mathbb{E}$ preserve the subobject classifiers, and are thus logical.

Once again, the fact that $\mathbb{S}\mathrm{ub}_{\mathbb{E}}$ sends finite (fibrewise) colimits to limits is straightforward from conditions (i) and (ii) of 3.3.9, so we have only to verify that, given $x\colon J \to I$ in $\mathcal{S}$ and $A \in \text{ob } \mathcal{E}^J$, we have $\mathrm{Sub}^I_{\mathbb{E}}(\Sigma_x A) \cong \Pi_x(\mathrm{Sub}^J_{\mathbb{E}}(A))$. For this, we note first that the functors Σ_x are conservative, by an easy application of 1.4.11. We claim next that they also preserve monomorphisms: for this, suppose we have $m\colon A' \rightarrowtail A$ in $\mathcal{E}^J$, and form the image factorization

$$\Sigma_x A' \xrightarrow{g} B \rightarrowtail^{f} \Sigma_x A$$

of $\Sigma_x m$. Now form the diagram

$$\begin{tikzcd} A' \arrow[r, "h"] \arrow[d, "\eta_{A'}"] & C \arrow[r, tail, "k"] \arrow[d] & A \arrow[d, "\eta_A"] \\ x^*\Sigma_x A' \arrow[r, "x^*g"] & x^*B \arrow[r, tail, "x^*f"] & x^*\Sigma_x A \end{tikzcd}$$

where the right-hand square is a pullback, and h is the unique factorization of $(m, (x^*g)\eta_{A'})$ through it. Then 1.4.10 tells us that $B \cong \Sigma_x C$, by an isomorphism which identifies g with $\Sigma_x h$. Since Σ_x is faithful, it reflects epimorphisms; so h is epic; but it is also monic since $kh = m$ is, and so it is an isomorphism. Thus g is an isomorphism, and $\Sigma_x m$ is monic.

Now we claim that we have a bijection between subobjects of A in $\mathcal{E}^J$ and subobjects of $\Sigma_x A$ in $\mathcal{E}^I$: the mapping in one direction is obtained by applying Σ_x to subobjects of A, and in the other direction we apply x^* to subobjects of $\Sigma_x A$ and pull back along η_A. Lemma 1.4.10 implies that the composite one way round is the identity, and the argument in the previous paragraph verifies that the composite the other way round is also the identity.

Finally, given an object K of $\mathcal{S}$ and a morphism $y\colon K \to I$, let us form the pullback

$$\begin{tikzcd} L \arrow[r, "z"] \arrow[d, "w"] & J \arrow[d, "x"] \\ K \arrow[r, "y"] & I \end{tikzcd} \quad ;$$

then morphisms $y \to \Pi_x(\mathrm{Sub}^J_{\mathbb{E}}(A))$ in $\mathcal{S}/I$ correspond to morphisms $z \to \mathrm{Sub}^J_{\mathbb{E}}(A)$ in $\mathcal{S}/J$, and hence to subobjects of z^*A in $\mathcal{E}^L$. By the argument just given, these correspond bijectively to subobjects of $\Sigma_w z^*A \cong y^*\Sigma_x A$ in $\mathcal{E}^K$, and hence to morphisms $y \to \mathrm{Sub}^I_{\mathbb{E}}(\Sigma_x A)$ in $\mathcal{S}/I$. It is straightforward to verify that these bijections are natural in y; and so we have an isomorphism $\Pi_x(\mathrm{Sub}^J_{\mathbb{E}}(A)) \cong \mathrm{Sub}^I_{\mathbb{E}}(\Sigma_x A)$, as required. □

We may now prove a result (the equivalence of (i) and (iv) below) which, in the case $\mathcal{S} = \mathbf{Set}$, reduces to something much more like Giraud's original theorem.

Theorem 3.3.11 *Let $\mathcal{S}$ be a topos, and $\mathbb{E}$ an $\mathcal{S}$-indexed category. The following conditions are equivalent:*

(i) *$\mathbb{E}$ is an $\mathcal{S}$-indexed ∞-pretopos with a separating family.*

(ii) *$\mathbb{E}$ is a locally small, cocomplete $\mathcal{S}$-indexed topos with a separating family.*

(iii) *There exists a bounded geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$ such that $\mathbb{E}$ is equivalent to the indexing of $\mathcal{E}$ over $\mathcal{S}$ induced by p^*.*

(iv) *There exists an internal category $\mathbb{C}$ in $\mathcal{S}$, such that $\mathbb{E}$ is equivalent to a reflective indexed subcategory of the indexed functor category $[\![\mathbb{C}, \mathbb{S}]\!]$ with cartesian reflector.*

Proof Combine 3.3.10, 3.1.2, 3.3.4 and A4.3.9. □

Suggestions for further reading: Chapman & Rowbottom [238], Diaconescu [280], Giraud [405].

B3.4 Colimits in $\mathfrak{Top}$

There is a sense in which this section does not belong in the present chapter. Half of it belongs in Chapter A4, since it consists of elementary results about geometric morphisms; and the other half is restricted to bounded toposes over a base, and so belongs in Chapter B4 (or even in Part C). However, the logical flow of ideas seems to point to this stage in the text as the correct one at which to gather together the results which follow.

We shall see in the cases which follow that, whenever a diagram D in $\mathfrak{Top}$ has a (possibly weighted) colimit, the underlying category of the colimit is always the (weighted) limit in $\mathfrak{CAT}$ of the corresponding diagram of toposes and inverse image functors. There is of course a reason for this: we have seen in 3.2.9 above that, for a topos $\mathcal{S}$ with a natural number object, the forgetful functor $U\colon (\mathfrak{Top}/\mathcal{S})^{\mathrm{op}} \to \mathfrak{CAT}$, which sends an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ to the underlying category of $\mathcal{E}$ and a geometric morphism to its inverse image functor, is representable, and so it must transform arbitrary weighted colimits to weighted limits. Hence, given a weighted diagram (D, W) in $\mathfrak{Top}$, provided there exists some topos $\mathcal{S}$ with a natural number object such that the diagram D can be regarded as living in $\mathfrak{Top}/\mathcal{S}$, then the weighted colimit of (D, W) if it exists will be the weighted limit of (UD^{op}, W) in $\mathfrak{CAT}$. (Note that if W is the trivial weighting, so that we are seeking an ordinary 'conical' colimit, then the colimit topos $\mathcal{L}$ itself, if it exists, will be a topos such that the diagram D lives in $\mathfrak{Top}/\mathcal{L}$; but this is not the case for more general weightings.) Since weighted limits are generally easy to construct in $\mathfrak{CAT}$, the major part of the work in showing that particular types of weighted colimits exist in $\mathfrak{Top}$ is thus to show that these weighted limits are in fact toposes.

We begin with the simplest case, that of finite coproducts.

Lemma 3.4.1 *The 2-category $\mathfrak{Top}$ has finite coproducts.*

Proof The terminal object of $\mathfrak{CAT}$, the category $\mathbf{1}$ with one object and one (identity) morphism, is initial in $\mathfrak{Top}$: for any topos $\mathcal{E}$, the unique functor $\mathcal{E} \to \mathbf{1}$ is the inverse image of a geometric morphism, whose direct image sends the unique object of $\mathbf{1}$ to the terminal object of $\mathcal{E}$, and this is clearly the only geometric morphism $\mathbf{1} \to \mathcal{E}$ up to unique 2-isomorphism. Similarly, if $\mathcal{E}$ and $\mathcal{F}$ are toposes, then the cartesian product $\mathcal{E} \times \mathcal{F}$ is a topos (we can regard it as a special case of the glueing construction of A2.1.12, where F is the constant functor with value 1), and the projections from it to $\mathcal{E}$ and $\mathcal{F}$ are inverse image functors (their right adjoints send objects A and B to $(A, 1)$ and $(1, B)$ respectively). And, given an arbitrary cospan

$$\mathcal{E} \xrightarrow{\ f\ } \mathcal{G} \xleftarrow{\ g\ } \mathcal{F}$$

in $\mathfrak{Top}$, we have an induced morphism $h\colon \mathcal{E} \times \mathcal{F} \to \mathcal{G}$ given by $h^*(C) = (f^*(C), g^*(C))$ and $h_*(A, B) = f_*(A) \times g_*(B)$; it is clear that the construction $(f, g) \mapsto h$ is functorial, and that, together with the operation of composing

with the morphisms whose inverse images are the projections, it defines an equivalence of categories between $\mathfrak{Top}(\mathcal{E} \times \mathcal{F}, \mathcal{G})$ and $\mathfrak{Top}(\mathcal{E}, \mathcal{G}) \times \mathfrak{Top}(\mathcal{F}, \mathcal{G})$. □

We note that the initial object of $\mathfrak{Top}$ is strict, since if we are given a geometric morphism $f\colon \mathcal{E} \to \mathbf{1}$ then f^* preserves the isomorphism $0 \cong 1$, and so $\mathcal{E}$ is degenerate. Also, coproducts are disjoint and stable under pullback; for we can identify $\mathcal{E}$ and $\mathcal{F}$ with the complementary open subtoposes $(\mathcal{E} \times \mathcal{F})/(1,0)$ and $(\mathcal{E} \times \mathcal{F})/(0,1)$ of the coproduct, and any geometric morphism $f\colon \mathcal{G} \to \mathcal{E} \times \mathcal{F}$ induces a corresponding decomposition $\mathcal{G} \simeq \mathcal{G}/f^*(1,0) \times \mathcal{G}/f^*(0,1)$ of its domain.

The only reason for restricting to finite coproducts in 3.4.1 is the need to form a product inside $\mathcal{G}$ in the definition of h_*; if we are working in a context where the toposes concerned all have arbitrary set-indexed products, for example in the 2-category $\mathfrak{BTop}/\mathbf{Set}$, then we may conclude by exactly similar arguments that an arbitrary set-indexed family of toposes has a coproduct, which is simply the product of their underlying categories. We shall generalize this observation in 3.4.7(i) below.

The glueing construction, or some variant of it, is used in most of the cases where colimits can be constructed in $\mathfrak{Top}$ by elementary means. A good example is the following:

Lemma 3.4.2 *The 2-category $\mathfrak{Top}$ has cocomma objects.*

Proof Let

$$\mathcal{F} \xleftarrow{\ f\ } \mathcal{E} \xrightarrow{\ g\ } \mathcal{G}$$

be a span in $\mathfrak{Top}$. To form the cocomma object $(f \uparrow g)$, we need to consider the comma object $(f^* \downarrow g^*)$ in $\mathfrak{CAT}$; but this is the usual comma category, whose objects are triples (B, C, h) where $B \in \text{ob}\, \mathcal{F}$, $C \in \text{ob}\, \mathcal{G}$ and $h\colon f^*(B) \to g^*(C)$. And since morphisms $f^*(B) \to g^*(C)$ correspond bijectively to morphisms $B \to f_*g^*(C)$, this category is isomorphic to the comma category $(1_{\mathcal{F}} \downarrow f_*g^*)$, i.e. to $\mathbf{Gl}(f_*g^*)$. Moreover, f_*g^* is cartesian, so by A2.1.12 $\mathbf{Gl}(f_*g^*)$ is a topos, and by A4.1.12 we have geometric morphisms

$$\mathcal{F} \xrightarrow{\ q\ } \mathbf{Gl}(f_*g^*) \xleftarrow{\ p\ } \mathcal{G};$$

and the canonical isomorphism $q^*p_* \cong f_*g^*$ of A4.1.12 transposes to yield a natural transformation $f^*q^* \to g^*p^*$, i.e. a geometric transformation $qf \to pg$. So we have constructed a weighted cone of the required form under (f, g) in $\mathfrak{Top}$.

Given any other such weighted cone

$$\begin{tikzcd} \mathcal{E} \arrow[r, "g"] \arrow[d, "f"'] \arrow[dr, phantom, "\overset{\alpha}{\to}"] & \mathcal{G} \arrow[d, "h"] \\ \mathcal{F} \arrow[r, "k"] & \mathcal{H} \end{tikzcd},$$

there exists (up to unique isomorphism) a unique functor $l^*: \mathcal{H} \to \mathbf{Gl}(f_*g^*)$ with $p^*l^* \cong h^*$ and $q^*l^* \cong k^*$, namely $l^*(D) = (h^*(D), k^*(D), \overline{\alpha}_D)$ where $\overline{\alpha}_D: k^*(D) \to f_*g^*h^*(D)$ is the transpose of α_D across the adjunction $(f^* \dashv f_*)$. Moreover, l^* is cartesian since h^* and k^* are, and it has a right adjoint l_* sending $(C, B, u: B \to f_*g^*(C))$ to the pullback of

$$\begin{tikzcd} & & k_*(B) \arrow[d, "k_*(u)"] \\ h_*(C) \arrow[r, "h_*(\eta_C)"] & h_*g_*g^*(C) \arrow[r, "\hat{\alpha}_{g^*(C)}"] & k_*f_*g^*(C) \end{tikzcd}$$

where η is the unit of $(g^* \dashv g_*)$ and $\hat{\alpha}$ is the mate of α. So we have a geometric morphism $l: \mathbf{Gl}(f_*g^*) \to \mathcal{H}$; the construction of l is clearly functorial in the data (h, k, α), and defines (the other half of) the required equivalence between $\mathfrak{Top}\,(\mathbf{Gl}(f_*g^*), \mathcal{H})$ and the appropriate category of weighted cones. $\square$

If we take $f = g = 1_{\mathcal{E}}$ in 3.4.2, we obtain the result mentioned before 3.2.14 that the functor category $[\mathbf{2}, \mathcal{E}]$ (the Sierpiński topos over $\mathcal{E}$, cf. 3.2.11) serves as the tensor $\mathbf{2} \otimes \mathcal{E}$ in $\mathfrak{Top}$. More generally, if $\mathcal{C}$ is any finite category, then the functor category $[\mathcal{C}, \mathcal{E}]$ (which is of course the cotensor $\mathcal{C} \pitchfork \mathcal{E}$ in $\mathfrak{CAT}$ – cf. 1.1.4(*b*)) can similarly be shown to have the universal property of a tensor $\mathcal{C} \otimes \mathcal{E}$ in $\mathfrak{Top}$. (Recall that we showed that $[\mathcal{C}, \mathcal{E}]$ was a topos in 2.3.18.)

The next example uses a generalization of the glueing construction.

Lemma 3.4.3 *Pushouts of pairs of inclusions exist in* $\mathfrak{Top}$.

Proof Let

$$\mathcal{F} \xleftarrow{\ f\ } \mathcal{E} \xrightarrow{\ g\ } \mathcal{G}$$

be a pair of inclusions with common domain. Consider the functor $G: \mathcal{F} \times \mathcal{G} \to \mathcal{F} \times \mathcal{G}$ which sends (B, C) to $(B \times f_*g^*(C), C \times g_*f^*(B))$; this functor is clearly cartesian, and it has a comonad structure whose counit is induced by the projections $B \times f_*g^*(C) \to B$ and $C \times g_*f^*(B) \to C$, and whose comultiplication is induced by

$$B \times f_*g^*(C) \xrightarrow{(\pi_1, \pi_2, \pi_2, \theta\pi_1)} B \times f_*g^*(C) \times f_*g^*(C) \times f_*g^*g_*f^*(B)$$

(where θ is the composite of the unit $\eta_B: B \to f_*f^*(B)$ and f_* applied to the inverse of the counit of $(g^* \dashv g_*)$) and the corresponding morphism with f and g interchanged. (The verification of the comonad identities is tedious but straightforward.) So the category of coalgebras $(\mathcal{F} \times \mathcal{G})_{\mathbb{G}}$ is a topos, by A4.2.1; but the objects of this topos can be identified with quadruples (B, C, α, β) where $\alpha: B \to f_*g^*(C)$ and $\beta: C \to g_*f^*(B)$ are morphisms satisfying coherence conditions equivalent to saying that their transposes $f^*(B) \to g^*(C)$ and

$g^*(C) \to f^*(B)$ are mutually inverse isomorphisms. So we may identify the topos $(\mathcal{F} \times \mathcal{G})_{\mathbb{G}}$ with the (pseudo-)pullback of f^* and g^* in $\mathfrak{CAT}$. As in 3.4.1 and 3.4.2, the verification that it has the universal property of a pushout in $\mathfrak{Top}$ is now straightforward. □

From the proof of 3.4.3, we may obtain a result reminiscent of A2.4.3: note that it includes the result, already mentioned, that coproducts in $\mathfrak{Top}$ are disjoint.

Scholium 3.4.4 *Let*

$$\begin{tikzcd} \mathcal{E} \arrow[r, "f"] \arrow[d, "g"'] & \mathcal{F} \arrow[d, "h"] \\ \mathcal{G} \arrow[r, "k"] & \mathcal{H} \end{tikzcd}$$

be a pushout square in $\mathfrak{Top}$ where f and g are inclusions. Then h and k are also inclusions, and the square is also a pullback.

Proof By 3.4.3, we may identify $\mathcal{H}$ with the category whose objects are triples (B, C, α) where $B \in \operatorname{ob} \mathcal{F}$, $C \in \operatorname{ob} \mathcal{G}$ and α is an isomorphism $f^*B \to g^*C$ in $\mathcal{E}$. The inverse images of h and k are the two projections; and an easy calculation shows that $h_*(B)$ may be taken to be the triple $(B, g_*f^*B, (\epsilon_{f^*B})^{-1})$, where ϵ is the counit of $(g^* \dashv g_*)$. Hence the counit of $(h^* \dashv h_*)$ is an isomorphism; and similarly for k. To show that the square is a pullback, we may identify $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$ with full subcategories of $\mathcal{H}$, consisting respectively of those objects (B, C, α) for which B and C are sheaves for the appropriate local operators on $\mathcal{F}$ and $\mathcal{G}$ (resp. C is a sheaf, B is a sheaf). It is clear that $\mathcal{E}$ is simply the intersection of $\mathcal{F}$ and $\mathcal{G}$; so the fact that the square is a pullback in $\mathfrak{Top}$ is immediate from the criterion of A4.3.11(iii) for a geometric morphism to factor through an inclusion. □

Corollary 3.4.5 *Any inclusion is a regular monomorphism (that is, the pseudo-limit of a coreflexive pair) in $\mathfrak{Top}$.*

Proof Set $f = g$ in 3.4.4, and combine it with A1.2.10. □

The reader should be warned not to read into 3.4.5 more than it says: it does *not* assert that if (h, k) is the cokernel-pair of an inclusion f, then f is the (pseudo-)equalizer of the pair (h, k), but rather that f is the limit of the diagram formed by h, k and the 'codiagonal' morphism d which is a common splitting for g and h. The point is that, when we are working in a 2-category, these two limits do not necessarily coincide: a cone over the diagram formed by h and k is a morphism g equipped with a 2-isomorphism $\alpha\colon hg \to kg$, whereas a cone over the diagram formed by h, k and d must satisfy the additional condition

that $d \circ \alpha$ is (modulo the coherence isomorphisms) the identity 2-cell on g. If the identity functor on the domain $\mathcal{E}$ of f admits nontrivial automorphisms (as is the case, for example, if $\mathcal{E} = [G, \mathbf{Set}]$ for some group G with nontrivial centre), then f will not have the required universal property among cones of the first type; but it will be universal among cones of the second type. We shall encounter a similar phenomenon when we discuss the difference between 1-dimensional and 2-dimensional colimits of simplicial objects in 3.4.11 below.

We do not know of any other cases where 'conical' colimits can be constructed in $\mathfrak{Top}$ by elementary means. However, as noted by G. C. Wraith [1235], a generalized version of the glueing construction suffices to construct (op)lax colimits, as defined in 1.1.6, for arbitrary finite (op)lax diagrams in $\mathfrak{Top}$. Recall that a lax diagram D of shape J in $\mathfrak{Top}$ assigns a topos $D(j)$ to each object j of J, a geometric morphism $D(\alpha)\colon D(j) \to D(j')$ to each morphism $\alpha\colon j \to j'$ of J, a geometric transformation $\phi_j\colon 1_{D(j)} \to D(1_j)$ to each object j, and a geometric transformation $\phi_{\alpha,\beta}\colon D(\beta)D(\alpha) \to D(\beta\alpha)$ to each composable pair (α, β), subject to appropriate coherence conditions. (In Section B1.1 we actually considered lax diagrams over 2-categories, but for simplicity we shall assume here that J has no 2-cells other than identities.)

Lemma 3.4.6 *Any finite lax (resp. oplax) diagram in $\mathfrak{Top}$ has a lax (resp. oplax) colimit.*

Proof We consider the lax case in detail. Suppose given a lax diagram $D\colon J \to \mathfrak{Top}$ as above; we form the topos $\mathcal{E} = \prod_{j \in J} D(j)$ (that is, the coproduct of the $D(j)$ in $\mathfrak{Top}$), and define a functor $G\colon \mathcal{E} \to \mathcal{E}$ by

$$G(A_j \mid j \in J) = \Big(\prod_{\alpha j' \to j} D(\alpha)_*(A_{j'}) \mid j \in J \Big) .$$

G is clearly a cartesian functor, since each $D(\alpha)_*$ is cartesian. We claim that G has a comonad structure: the counit $\epsilon\colon G(A_j \mid j \in J) \to (A_j \mid j \in J)$ is the morphism whose jth component is the composite

$$\prod_{\alpha} D(\alpha)_*(A_{j'}) \xrightarrow{\pi_{1_j}} D(1_j)_*(A_j) \xrightarrow{\overline{\phi_j}} A_j$$

where $\overline{\phi_j}$ is the mate of ϕ_j. Similarly, the comultiplication δ is the unique morphism making

$$\begin{tikzcd}
\prod_{\alpha} D(\alpha)_*(A_{j'}) \arrow[r, "\delta"] \arrow[d, "\pi_{\beta\alpha}"] & \prod_{\alpha,\beta} D(\beta)_* D(\alpha)_*(A_{j''}) \arrow[d, "\pi_{\alpha,\beta}"] \\
D(\beta\alpha)_*(A_{j''}) \arrow[r, "\overline{\phi_{\alpha,\beta}}"] & D(\beta)_* D(\alpha)_*(A_{j''})
\end{tikzcd}$$

commute for each composable pair (α, β). The comonad identities for ϵ and δ follow straightforwardly from the coherence conditions on the ϕ's.

Hence by A4.2.1 we have a topos $\mathcal{E}_{\mathbb{G}}$ of coalgebras for the comonad $\mathbb{G} = (G, \epsilon, \delta)$. We claim that this is (the vertex of) a lax colimit for the given lax diagram in $\mathfrak{Top}$. For giving an object $(A_j \mid j \in J)$ of $\mathcal{E}$ a $\mathbb{G}$-coalgebra structure is easily verified to be equivalent to specifying a family of morphisms $\overline{\theta_\alpha}: A_j \to D(\alpha)_*(A_{j'})$ for each $\alpha: j' \to j$ in J, satisfying appropriate compatibility conditions with the $\overline{\phi_j}$ and $\overline{\phi_{\alpha,\beta}}$; hence it is also equivalent to specifying a family of morphisms $\theta_\alpha : D(\alpha)^*(A_j) \to A_{j'}$ satisfying compatibility conditions with the ϕ_j and $\phi_{\alpha,\beta}$. Thus we see that the projections $\mathcal{E}_{\mathbb{G}} \to D(j)$ are the legs of a lax limit cone over the lax diagram in $\mathfrak{CAT}$ formed by the $D(\alpha)^*$; but they are inverse image functors, and as usual it is easy to see that the corresponding lax cone under the diagram in $\mathfrak{Top}$ is a lax colimit.

The oplax case is similar, except that we use the inverse image functors $D(\alpha)^*$ to define our comonad $\mathbb{G}$. □

Of course, an ordinary finite diagram may be considered as a lax (or oplax) diagram in which the ϕ_j and $\phi_{\alpha,\beta}$ are isomorphisms. In particular, considering a single geometric morphism $f: \mathcal{F} \to \mathcal{E}$ as a diagram of shape $\mathbf{2}$ in $\mathfrak{Top}$, we note that its lax and oplax colimits are the cocomma objects $(f \uparrow 1_{\mathcal{F}})$ and $(1_{\mathcal{F}} \uparrow f)$ respectively. It is easy to see that the construction of 3.4.6 reduces to that of 3.4.2 in these two particular cases.

Now we turn from $\mathfrak{Top}$ to $\mathfrak{Top}/\mathcal{S}$. We note that the latter is itself enriched, not just over $\mathfrak{CAT}$, but over the 2-category $\mathfrak{CAT}_{\mathcal{S}}$ of $\mathcal{S}$-indexed categories: given $\mathcal{S}$-toposes $p: \mathcal{E} \to \mathcal{S}$ and $q: \mathcal{F} \to \mathcal{S}$ and an object I of $\mathcal{S}$, we define an 'I-indexed family of geometric morphisms $\mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$' to be a geometric morphism $\mathcal{F}/q^*I \to \mathcal{E}/p^*I$ over $\mathcal{S}/I$, with transition functors induced by pulling back along geometric morphisms of the form $\mathcal{S}/J \to \mathcal{S}/I$. (In fact $\mathfrak{BTop}/\mathcal{S}$ is even enriched over $\mathfrak{Cat}_{\mathcal{S}}$, the 2-category of locally internal categories over $\mathcal{S}$: as we shall see in 4.1.6 below, if $\mathcal{E}$ is bounded over $\mathcal{S}$, then there exists an object E of $\mathcal{E}$ such that any geometric transformation $\alpha: f \to g$ between geometric morphisms $f, g: \mathcal{F} \rightrightarrows \mathcal{E}$ over $\mathcal{S}$ is uniquely determined by its component at E, and hence we can construct an object of $\mathcal{S}$ indexing the geometric transformations between f and g. However, we shall not need this extra structure.)

In considering colimits in $\mathfrak{BTop}/\mathcal{S}$, it therefore makes sense to ask for weighted colimits of diagrams indexed by internal categories in $\mathcal{S}$, where the weighting functor also takes values in $\mathfrak{Cat}(\mathcal{S})$. We shall not go into the general theory of such weighted colimits, though in fact it is very similar to the case $\mathcal{S} = \mathbf{Set}$; in particular, the analogue of Theorem 1.1.5 still holds. But we note a couple of special cases of the notion.

Lemma 3.4.7 *Let $\mathcal{S}$ be a topos.*

(i) *$\mathfrak{Top}/\mathcal{S}$ and $\mathfrak{BTop}/\mathcal{S}$ both have $\mathcal{S}$-indexed coproducts.*

(ii) *$\mathfrak{Top}/\mathcal{S}$ and $\mathfrak{BTop}/\mathcal{S}$ both have tensors with internal categories in $\mathcal{S}$.*

Proof (i) Clearly, an I-indexed family of $\mathcal{S}$-toposes is just a topos $q\colon \mathcal{F} \to \mathcal{S}/I$ over $\mathcal{S}/I$. To form its coproduct, we simply form the composite $\mathcal{F} \to \mathcal{S}/I \to \mathcal{S}$; for, given any $p\colon \mathcal{E} \to \mathcal{S}$, morphisms $\mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ correspond (by a particular case of 3.2.12) to morphisms $\mathcal{F} \to \mathcal{E}/p^*I$ over $\mathcal{S}/I$. And the fact that $\mathfrak{BTop}/\mathcal{S}$ is closed under $\mathcal{S}$-indexed coproducts is a particular case of 3.1.10(i).

(ii) In the same spirit, given $q\colon \mathcal{F} \to \mathcal{S}$ and an internal category $\mathbb{C}$ in $\mathcal{S}$, we define the tensor $\mathbb{C} \otimes \mathcal{F}$ in $\mathfrak{Top}/\mathcal{S}$ to be the diagram category $[q^*\mathbb{C}, \mathcal{F}]$. Clearly, this is the cotensor $\mathbb{C} \pitchfork \mathcal{F}$ in the 2-category of $\mathcal{S}$-indexed categories, by 2.3.13; so it is easy to see that if we are given a diagram of shape $\mathbb{C}$ in $\mathfrak{Top}/\mathcal{S}\,(\mathcal{F}, \mathcal{E})$ (that is, a geometric morphism $f\colon \mathcal{F}/q^*C_0 \to \mathcal{E}/p^*C_0$ over $\mathcal{S}/C_0$, equipped with an action transformation $d_1^*(f) \to d_0^*(f)$ in $\mathfrak{Top}/(\mathcal{S}/C_1)$ satisfying the appropriate equations), then we obtain an indexed functor $g^*\colon \mathcal{E} \to [\mathbb{C}, \mathcal{F}]$: given an object A of $\mathcal{E}$, we define g^*A to be the object $f^*(C_0^*(A))$ of $\mathcal{F}/C_0$, equipped with the component at A of the given geometric transformation. Defining the direct image functor g_* is a little more tricky: we need to introduce an auxiliary internal category $\mathbb{D}$, as follows. Its object of objects D_0 is C_1, and its object of morphisms D_1 is the colimit of the diagram

$$\begin{tikzcd} & C_0 \arrow[dl, "s_0"'] \arrow[d, "s_0"] \arrow[dr, "s_0"] & \\ C_1 & C_1 & C_1 \end{tikzcd}$$

with the domain map $d_1\colon D_1 \to D_0$ being induced by the triple $(s_0 d_1, 1_{C_1}, s_0 d_0)$ of morphisms $C_1 \to C_1$ and the codomain map induced by $(1_{C_1}, 1_{C_1}, 1_{C_1})$; the inclusion of identities $D_0 \to D_1$ is the middle leg of the colimit cone defining D_1. (Composition is easy to define, since a straightforward diagram-chase shows that at least one member of any composable pair in $\mathbb{D}$ is an identity morphism.) Now, given an object $(F \to C_0, \phi\colon d_1^*F \to d_0^*F)$ of $[\mathbb{C}, \mathcal{F}]$ (we shall suppress all mention of the functors p^* and q^* from now on, to simplify the notation), we may construct a diagram of shape $\mathbb{D}$ in $\mathcal{E}$: we form the object $(d_1^*g)_*(d_0^*F)$ of $\mathcal{E}/C_1 = \mathcal{E}/D_0$, equipped with the D_1-action in which elements of D_1 belonging to the first copy of C_1 in D_1 act via $(d_1^*g)_*(\phi)$ and those in the third copy of C_1 act via the $\mathbb{C}$-action on f_* (which is of course contravariant, since $\mathbb{C}$ acts covariantly on f^*). Now we define $g^*(F, \phi)$ to be the limit in $\mathcal{E}$ of this internal diagram. Checking that this does indeed define a right adjoint to g^* requires a lot of tedious verification, but it should be clear enough if one decodes the above definition in the case $\mathcal{S} = \mathbf{Set}$ (and compares it with the construction we gave for the direct image functor l_* in the proof of 3.4.2).

In the opposite direction, we have a canonical geometric morphism $\mathcal{F}/C_0 \to [\mathbb{C}, \mathcal{F}]$ over $\mathcal{S}$ (indeed, over $\mathcal{F}$) whose inverse image is the forgetful functor; as in part (i), we may equivalently regard this as a morphism $\mathcal{F}/C_0 \to [\mathbb{C}, \mathcal{F}]/C_0$ over $\mathcal{S}/C_0$, i.e. a C_0-indexed family of geometric morphisms $\mathcal{F} \to [\mathbb{C}, \mathcal{F}]$ over $\mathcal{S}$; and

this clearly admits a $\mathbb{C}$-action map, i.e. we can regard it as a dagram of shape $\mathbb{C}$ in $\mathfrak{Top}/\mathcal{S}\,(\mathcal{F},[\mathbb{C},\mathcal{F}])$. So composition with this diagram yields a functor from $\mathfrak{Top}/\mathcal{S}\,([\mathbb{C},\mathcal{F}],\mathcal{E})$ to $[\mathbb{C},\mathfrak{Top}/\mathcal{S}\,(\mathcal{F},\mathcal{E})]$; it is straightforward to verify that this is inverse (up to natural isomorphism) to the construction defined above.

As in part (i), the 'bounded' case of the result follows immediately from the case already done, plus the fact that $[q^*\mathbb{C},\mathcal{F}]$ is bounded over $\mathcal{F}$ and hence over $\mathcal{S}$. □

Of course, in the bounded case of (ii), we do not have to go through the hard work of defining g_*: we may simply observe that g^* is cocontinuous as an $\mathcal{S}$-indexed functor, and then use 2.4.6 to deduce that it has a right adjoint.

Whilst we are on the subject of $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ as an $\mathcal{S}$-indexed category, we note the following result, which can be viewed as an extension of A4.1.13. (Recall that splittings of idempotents in a category are essentially the same thing as colimits indexed by finite filtered categories.)

Lemma 3.4.8 *For any two bounded $\mathcal{S}$-toposes $\mathcal{E}$ and $\mathcal{F}$, the $\mathcal{S}$-indexed category $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ has colimits indexed by filtered $\mathcal{S}$-indexed categories.*

Proof Given a diagram of geometric morphisms $f_c\colon \mathcal{F}\to\mathcal{E}$ indexed by the objects of a filtered category $\mathbb{C}$ in $\mathcal{S}$, we may form the 'pointwise' colimit f^*_∞ of the functors f^*_c in the $\mathcal{S}$-indexed category of $\mathcal{S}$-indexed functors $\mathbb{E}\to\mathbb{F}$. The latter is cartesian by 2.6.8; but it is also clearly $\mathcal{S}$-cocontinuous; so it has a right adjoint $f_{\infty *}$ by 2.4.6. It is now routine to verify that the adjoint pair $(f^*_\infty \dashv f_{\infty *})$ defines a geometric morphism $\mathcal{F}\to\mathcal{E}$ over $\mathcal{S}$, and that it is the colimit of the f_c in $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$. □

We now revert to the problem of constructing colimits in $\mathfrak{BTop}/\mathcal{S}$ itself, rather than in its hom-categories. In order to do this in more complicated cases, we shall have to rely on Giraud's theorem: specifically, on the form of it proved in 3.3.11, where we saw that bounded $\mathcal{S}$-toposes are the same thing as $\mathcal{S}$-indexed ∞-pretoposes with separating families. The point is that the inverse image of a geometric morphism over $\mathcal{S}$ is a morphism of $\mathcal{S}$-indexed ∞-pretoposes (in the sense that it preserves all the structure involved in the definition of such categories), and it is easy to see that any $\mathcal{S}$-indexed weighted limit of $\mathcal{S}$-indexed ∞-pretoposes (and functors preserving this structure) is an $\mathcal{S}$-indexed ∞-pretopos. Since (as already indicated in this section) we expect weighted colimits of diagrams of toposes to be weighted limits of their underlying categories and inverse image functors, it therefore only remains to show that the limit in question has a separating family in order to deduce that it is a bounded $\mathcal{S}$-topos, and hence is the desired colimit in $\mathfrak{BTop}/\mathcal{S}$.

However, even for the simple case of coequalizers, showing that the equalizer in $\mathfrak{Cat}_{\mathcal{S}}$ has a separating family is by no means simple. From now on, we shall need to assume that $\mathcal{S}$ has a natural number object; in fact, for notational convenience, we shall treat it as if it were **Set**, but the reader who has followed the

material of Part B thus far should have little conceptual difficulty (as opposed to the practical difficulty of actually writing it down!) in translating the argument which follows into '$\mathcal{S}$-indexed language'. The argument which we give is essentially that of I. Moerdijk [831], but similar arguments were given independently by M. Tierney and P. Freyd.

So let us suppose given a parallel pair $f, g\colon \mathcal{F} \rightrightarrows \mathcal{E}$ of geometric morphisms over $\mathcal{S}$. Of course, when we refer to the equalizer of f^* and g^* in $\mathfrak{Cat}_{\mathcal{S}}$, we mean their pseudo-equalizer; that is, the category $\mathcal{G}$ whose objects are pairs (A, α) where $A \in \mathrm{ob}\,\mathcal{E}$ and α is an isomorphism $f^*A \to g^*A$ in $\mathcal{F}$. Morphisms $(A, \alpha) \to (B, \beta)$ in $\mathcal{G}$ are morphisms $A \to B$ in $\mathcal{E}$ commuting with α and β in the obvious sense; note in particular that the forgetful functor $\mathcal{G} \to \mathcal{E}$ is faithful (so that, once we have proved that $\mathcal{G}$ is a topos, we shall know that the coequalizer $\mathcal{E} \to \mathcal{G}$ is a surjection). Note also that the inverse image functor $p^*\colon \mathcal{S} \to \mathcal{E}$ has a canonical lifting to a functor $r^*\colon \mathcal{S} \to \mathcal{G}$, since we are given natural isomorphisms $f^*p^* \cong q^* \cong g^*p^*$; and the indexing of $\mathcal{G}$ over $\mathcal{S}$ which it acquires via this functor coincides with that which it acquires as the (pseudo-)equalizer of a pair of $\mathcal{S}$-indexed functors.

Now let $\mathbb{C}$ be an internal (to $\mathcal{S}$) full subcategory of $\mathbb{E}$ containing a separating family for the latter, so that $\mathcal{E}$ may be represented as a subtopos of $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$. For convenience, we shall suppose that $\mathbb{C}$ is closed under finite limits in $\mathbb{E}$ (we may clearly do this, since $\mathcal{S}$ has a natural number object). Borrowing the notation to be introduced formally in Section C2.4, we shall write $\mathcal{E} \simeq \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J)$, the topos of '$\mathcal{S}$-valued sheaves' on the internal site $(\mathbb{C}, J)$. We similarly choose an internal site $(\mathbb{D}, K)$ for $\mathcal{F}$ over $\mathcal{S}$; but we require that, in addition to a separating family for $\mathcal{F}$ over $\mathcal{S}$, the objects of $\mathbb{D}$ should include also the images under both f^* and g^* of the objects of $\mathbb{C}$. Thus f^* and g^* restrict to internal functors $F, G\colon \mathbb{C} \rightrightarrows \mathbb{D}$, which are cartesian and cover-preserving (that is, they are morphisms of sites in the sense of C2.3.1); and they induce the geometric morphisms f and g in the manner of C2.3.4. Explicitly, this means that f^* may be computed in terms of sheaves by first forming the left Kan extension along F (which yields only a functor $\mathbb{D}^{\mathrm{op}} \to \mathcal{S}$, not a sheaf in general) and then applying the associated K-sheaf functor (cf. C2.2.6).

Next we proceed to enlarge $\mathbb{C}$: let $\mathbb{C}^{(0)} = \mathbb{C}$ and, for $n \geq 1$, define $\mathbb{C}^{(n)}$ to be the internal full subcategory of $\mathcal{E}$ whose objects are coproducts $\coprod_{i \in I^+} A_i$ of objects of $\mathbb{C}^{(n)}$, where $I^+ = I \amalg 1$ is the set obtained by adding a singleton to the index set of some cover $(B_i \to B \mid i \in I)$ in the coverage K on $\mathbb{D}$. Let $\mathbb{C}^{(\infty)}$ similarly consist of all countable coproducts $\coprod_{n \geq 0} A_n$, where A_n is an object of $\mathbb{C}^{(n)}$ for each n, and let $\overline{\mathbb{C}}$ be the full subcategory on all quotients in $\mathcal{E}$ of objects of $\mathbb{C}^{(\infty)}$. It is clear that all these families of objects can be indexed by appropriate objects of the base topos $\mathcal{S}$, so $\overline{\mathbb{C}}$ is indeed an internal category.

We claim that the indexed family of all pairs (A, α), where A is an object of $\overline{\mathbb{C}}$ and α is an isomorphism $f^*(A) \cong g^*(A)$ in $\mathcal{F}$, is a separating family for $\mathbb{G}$. To prove this, let $u, v\colon (C, \gamma) \rightrightarrows (D, \delta)$ be two morphisms in $\mathcal{G}$ such that $uw = vw$

for every $w\colon (A,\alpha) \to (C,\gamma)$ with A an object of $\overline{\mathbb{C}}$; we must show that $u = v$. Thinking of C and D as sheaves on the site $(\mathbb{C}, J)$, let x be an element of $C(A)$ for some $A \in \mathrm{ob}\,\mathbb{C}$ (which we may equivalently consider as a morphism $A \to C$ in $\mathcal{E}$). We shall show that there is a subobject $\overline{C} \rightarrowtail C$ containing the element x, which is the image of a morphism $A \to C$ with $A \in \mathrm{ob}\,\mathbb{C}^{(\infty)}$ (and is therefore an object of $\overline{\mathbb{C}}$), such that γ restricts to an isomorphism $f^*(\overline{C}) \to g^*(\overline{C})$; then we can take w to be the inclusion $\overline{C} \rightarrowtail C$, and deduce that u and v agree on the element x.

Now x clearly determines a morphism $f^*(x)\colon f^*A \to f^*C$, or equivalently an element of $f^*(C)(FA)$, corresponding to the element of the left Kan extension of C determined by the element x and the identity morphism $FA \to FA$. Applying γ_{FA} to this, we get an element of $g^*(C)(FA)$: this may be represented by a K-covering family $(B_i \to FA \mid i \in I)$ and a family of elements y_i of the left Kan extension of C along G; and the y_i may in turn be represented by morphisms $(B_i \to GA_i)$ and elements $z_i \in C(A_i)$. Let A_1 be the coproduct of the family obtained by adjoining A itself to the A_i, $i \in I$, and $x_1 : A_1 \to C$ the morphism induced by the z_i; then the construction ensures that the composite

$$f^*(A) \xrightarrow{f^*(x)} f^*(C) \xrightarrow{\gamma_C} g^*(C)$$

factors through the image of $g^*(x_1)\colon g^*(A_1) \to g^*C$.

We now proceed recursively to define a sequence of objects A_n of $\mathbb{C}^{(n)}$ and morphisms $x_n\colon A_n \to C$ in $\mathcal{E}$, such that each A_{n+1} contains A_n as a subobject, and x_n is the restriction of x_{n+1} to this subobject. For even values of n we choose x_{n+1}, as above, so that the composite $\gamma_C \cdot f^*(x_n)$ factors through the image of $g^*(x_{n+1})$, and for odd n we do the same thing with f and g interchanged (and γ replaced by its inverse). Let A_∞ be the coproduct of the A_n, and $\overline{C} \rightarrowtail C$ the image of the morphism $A_\infty \to C$ induced by the x_n. Then $\overline{C}$ is the union of the increasing sequence of subobjects C_n which are the images of the x_n; and since γ_C maps $f^*(C_{2n})$ into $g^*(C_{2n+1})$ for all n, and its inverse maps $g^*(C_{2n+1})$ into $f^*(C_{2n+2})$, it is clear that γ_C must restrict to an isomorphism $f^*(\overline{C}) \to g^*(\overline{C})$, as required.

We have thus done all the hard work in the proof of

Theorem 3.4.9 *The 2-category $\mathfrak{BTop}/\mathcal{S}$ has coequalizers, for any topos $\mathcal{S}$ with a natural number object.*

Proof As above, given $f, g\colon \mathcal{F} \rightrightarrows \mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$, we form the (pseudo-)equalizer $\mathbb{G}$ of f^* and g^* in $\mathfrak{Cat}_{\mathcal{S}}$. The verification that $\mathbb{G}$ is a $\mathcal{S}$-indexed ∞-pretopos, as defined in 3.3.9, is straightforward, since $\mathbb{E}$ is one, and the functors f^* and g^* preserve all the structure involved in the definition. The argument above shows that $\mathbb{G}$ has an $\mathcal{S}$-indexed separating family; so by 3.3.10 it is an $\mathcal{S}$-indexed topos, i.e. it corresponds to a (bounded) $\mathcal{S}$-topos $(r\colon \mathcal{G} \to \mathcal{S})$. Moreover, the forgetful functor $\mathbb{G} \to \mathbb{E}$ preserves all $\mathcal{S}$-indexed colimits; so it has a right adjoint by 2.4.6, and is thus (the indexed version of) the inverse image of a

geometric morphism $h: \mathcal{E} \to \mathcal{G}$ over $\mathcal{S}$. The fact that $hf \cong hg$ is clear from the construction; but if we are given any geometric morphism $k: \mathcal{E} \to \mathcal{H}$ over $\mathcal{S}$ satisfying $kf \cong kg$, then k^* certainly factors through h^*, and since we observed earlier that h is surjective we can use the adjoint lifting theorem A1.1.3 to factor the right adjoint k_* through h_*, and thus obtain a geometric morphism $\mathcal{G} \to \mathcal{H}$ over $\mathcal{S}$. So $h: \mathcal{E} \to \mathcal{G}$ has the universal property of a coequalizer in $\mathfrak{BTop}/\mathcal{S}$. □

As suggested earlier, we may combine 3.4.7 and 3.4.9 to obtain arbitrary $\mathcal{S}$-indexed weighted colimits of bounded $\mathcal{S}$-toposes, where $\mathcal{S}$ has a natural number object. We shall not consider these in full generality, but we pause to note one particular case:

Remark 3.4.10 We may construct coinserters in $\mathfrak{BTop}/\mathcal{S}$ (cf. 1.1.4(d)) by a construction very like that of coequalizers. That is, to form the coinserter of $f, g: \mathcal{F} \rightrightarrows \mathcal{E}$, we form the inserter of their inverse image functors (that is, the category whose objects are pairs (A, α) where $A \in \text{ob}\,\mathcal{E}$ and $\alpha: f^*A \to g^*A$ in $\mathcal{F}$ is a not-necessarily-invertible morphism), and we may prove as in 3.4.9 that this category is an $\mathcal{S}$-indexed ∞-pretopos with a separating family. As previously, the only hard part is constructing a separating family; and we may do this in the same way as we did for coequalizers – except that the proof that the construction works is slightly easier, because we do not have to 'work from both sides' at the inductive stage of defining the objects A_n.

Without going into further technicalities about $\mathcal{S}$-indexed weighted colimits, we note that 3.4.7(i) and 3.4.9 do suffice to construct arbitrary '$\mathcal{S}$-indexed conical colimits', that is colimits of functors $\mathbb{C} \to \mathfrak{BTop}/\mathcal{S}$, where $\mathbb{C}$ is an internal category in $\mathcal{S}$, using the usual construction of colimits via coproducts and coequalizers.

Next, we need to say a bit about the particular case of colimits of simplicial toposes, in preparation for our work on descent morphisms in $\mathfrak{BTop}/\mathcal{S}$ in Chapter C5. Of course, by a *simplicial topos* we simply mean a (pseudo-)functor $\mathbf{\Delta}^{\text{op}} \to \mathfrak{Top}$, where $\mathbf{\Delta}$ is the category of nonzero finite ordinals and order-preserving maps between them, as in Section B2.3 above. We shall use the standard notation for simplicial objects: if $\mathcal{E}_\bullet$ is a simplicial topos, $\mathcal{E}_n$ will denote the image of the ordinal $n+1$ under the functor, and $d_i^n: \mathcal{E}_n \to \mathcal{E}_{n-1}$ (resp. $s_j^{n-1}: \mathcal{E}_{n-1} \to \mathcal{E}_n$) the images of the standard injections $n \to (n+1)$ (resp. surjections $(n+1) \to n$). If $\mathcal{S}$ has a natural number object, then it is clear that $\mathbf{\Delta}$ may be identified with an internal category in $\mathcal{S}$, either by the method of 2.3.12(b) or by explicitly constructing it as a subcategory of the internal category $\mathbb{S}_f$ of 3.2.9. So if we are given a diagram of shape $\mathbf{\Delta}^{\text{op}}$ in $\mathfrak{BTop}/\mathcal{S}$ we may compute its colimit as suggested above.

When we are working in an ordinary 1-dimensional category, the colimit of a simplicial object $A_\bullet$ is just the coequalizer of $d_0, d_1: A_1 \rightrightarrows A_0$, since the inclusion $\mathbf{\Delta}_1 \to \mathbf{\Delta}$ is initial (cf. 2.5.12), where $\mathbf{\Delta}_1$ is the subcategory of $\mathbf{\Delta}$ consisting

of the ordinals 1 and 2 and the two injections $1 \rightrightarrows 2$. But for diagrams in a 2-category, where we have to take the coherence 2-isomorphisms into account, we need to retain a bit more of the structure of $\mathbf{\Delta}$. Let $\mathbf{\Delta}_2$ be the subcategory of $\mathbf{\Delta}$ consisting of the ordinals 1, 2 and 3, the injections $1 \rightrightarrows 2$, the surjection $2 \to 1$, the three injections $2 \to 3$ and all composites of these morphisms. We shall refer to a diagram of shape $\mathbf{\Delta}_2^{\mathrm{op}}$ as a '2-truncated simplicial object'.

Lemma 3.4.11 *The inclusion $\mathbf{\Delta}_2 \to \mathbf{\Delta}$ is* 2-initial; *that is, for any simplicial object $A_\bullet \colon \Delta^{\mathrm{op}} \to \mathfrak{K}$ in a 2-category $\mathfrak{K}$ (admitting the appropriate colimits), the canonical morphism from the colimit of the restriction of $A_\bullet$ to $\mathbf{\Delta}_2$, to the colimit of $A_\bullet$, is an equivalence.*

Proof Let us write A_* for the restriction of $A_\bullet$ to $\mathbf{\Delta}_2$. We shall show that any (pseudo-)cone under A_* in $\mathfrak{K}$ may be extended, uniquely up to unique isomorphism, to a cone under $A_\bullet$. By definition, a cone under A_* (with vertex B, say) consists of three 1-cells $A_0 \to B$, $A_1 \to B$ and $A_2 \to B$, plus six 2-isomorphisms corresponding to the six generating morphisms of $\mathbf{\Delta}_2$. But since the coherence isomorphisms tell us that $f_1 \cong f_0 d_0^1$ and $f_2 \cong f_0 d_0^1 d_0^2$, we may reduce these data to a single 1-cell $f \colon A_0 \to B$ plus a 2-isomorphism $\alpha \colon f d_1^1 \to f d_0^1$ satisfying the 'unit condition' that $\alpha \circ s_0^0$ is the composite of the isomorphisms $f d_1^1 s_0^0 \cong f \cong f d_0^1 s_0^0$ arising from the equations $d_1^1 s_0^0 = 1 = d_0^1 s_0^0$ in $\mathbf{\Delta}^{\mathrm{op}}$, and the 'cocycle condition' that $\alpha \circ d_1^2$ is (modulo the appropriate coherence isomorphisms) the vertical composite of $\alpha \circ d_2^2$ and $\alpha \circ d_0^2$. (These conditions are of course reminiscent of those in the definition of descent data for a sieve generated by a single morphism; cf. 1.5.5. However, it will be noted that we cannot omit the requirement that α be invertible here, as we did in 1.5.5, since $\mathbf{\Delta}$ contains only order-preserving maps and we do not have the 'twist isomorphism' $A_1 \to A_1$ available. See 3.4.14(c) below for further discussion of this point.)

Given such data, we may define a pseudo-cone under $A_\bullet$, as follows: $f_n \colon A_n \to B$ is the composite $f d_0^1 d_0^2 \cdots d_0^n$. For a morphism d_i^n, the coherence isomorphism $f_{n-1} d_i^n \cong f_n$ is the identity if $i = 0$, the coherence isomorphism arising from the equality $d_0^1 \cdots d_0^{n-1} d_i^n = d_0^1 \cdots d_0^{n-1} d_0^n$ in $\mathbf{\Delta}^{\mathrm{op}}$ if $1 \leq i \leq n-1$, and the composite of the coherence isomorphism arising from $d_0^1 \cdots d_0^{n-1} d_n^n = d_1^1 d_0^2 \cdots d_0^{n-1} d_0^n$ with $\alpha \circ d_0^2 \cdots d_0^n$ if $i = n$. Similarly, the isomorphism $f_{n+1} s_i^n \cong f_n$ is just a coherence isomorphism arising from an equality in $\mathbf{\Delta}^{\mathrm{op}}$, for all i. It is straightforward but tedious to verify that these isomorphisms satisfy the appropriate compatibility conditions for all the simplicial identities, so that we do indeed have a pseudo-cone. Conversely, given any pseudo-cone $(g_n \colon A_n \to B \mid n \in \mathbb{N})$ under $A_\bullet$, the coherence isomorphisms yield a natural isomorphism from it to the cone constructed as above by extending its restriction to A_*. And the operations of extension and restriction are also well-defined on 2-cells between cones. So we conclude that $\mathbf{Cone}(A_\bullet, -)$ and $\mathbf{Cone}(A_*, -)$ are naturally equivalent $\mathfrak{CAT}$-valued pseudofunctors on $\mathfrak{K}$; hence their representing objects, if they exist, are equivalent. □

Corollary 3.4.12 *Let $\mathcal{S}$ be a topos with a natural number object. Given a simplicial bounded $\mathcal{S}$-topos $\mathcal{E}_\bullet$, its colimit in $\mathfrak{BTop}/\mathcal{S}$ may be identified with the category* $\mathbf{Desc}(\mathcal{E}_\bullet)$ *whose objects are pairs (A, α) where $A \in \mathrm{ob}\,\mathcal{E}_0$ and $\alpha \colon d_1^* A \to d_0^* A$ is an isomorphism in $\mathcal{E}_1$ satisfying the unit condition that $s_0^*(\alpha)$ is (modulo coherence isomorphisms) the identity isomorphism $A \to A$, and the cocycle condition that $d_1^*(\alpha)$ is (isomorphic to) the composite of $d_2^*(\alpha)$ and $d_0^*(\alpha)$ in $\mathcal{E}_2$. In particular,* $\mathbf{Desc}(\mathcal{E}_\bullet)$ *is a bounded $\mathcal{S}$-topos.*

Proof If we construct the colimit of the restriction $\mathcal{E}_*$ of $\mathcal{E}_\bullet$ to $\mathbf{\Delta}_2^{\mathrm{op}}$, and simplify its description in the same way that we simplified the description of cones under $\mathbf{\Delta}_2^{\mathrm{op}}$ above, we arrive at the description of the colimit given in the statement. So the result follows from 3.4.11. □

Remark 3.4.13 A direct proof that $\mathbf{Desc}(\mathcal{E}_\bullet)$ is a bounded $\mathcal{S}$-topos may be given along exactly the same lines as 3.4.9: if we consider the full subcategory of the coequalizer of d_0 and d_1 consisting of those (A, α) for which α satisfies the unit and cocycle conditions, it is still an $\mathcal{S}$-indexed ∞-pretopos (being closed in the coequalizer under the appropriate operations), and the argument we gave to show that the coequalizer has a separating family goes through without change for the descent category. (Note that the descent category is closed under subobjects as a subcategory of the coequalizer; in fact the inclusion is the inverse image of a hyperconnected geometric morphism.)

Examples 3.4.14 (*a*) Let $\mathbb{C}$ be an internal category in a topos $\mathcal{E}$. As we saw in Section B2.3, we may regard $\mathbb{C}$ as a simplicial object $C_\bullet$ in $\mathcal{E}$; and since the assignment $A \mapsto \mathcal{E}/A$ is functorial by A4.1.2, this in turn gives rise to a simplicial bounded $\mathcal{E}$-topos $\mathcal{E}/C_\bullet$. If we form the topos $\mathbf{Desc}(\mathcal{E}/C_\bullet)$, we do not, as might have been expected, simply get the diagram topos $[\mathbb{C}, \mathcal{E}]$, because of the requirement, in the definition of an object (A, α) of this topos, that the structure map α should be invertible. Thus what we actually obtain is the topos $[\mathbb{G}, \mathcal{E}]$ where $\mathbb{G}$ is the groupoid reflection of $\mathbb{C}$, i.e. the category obtained by freely adjoining inverses for all morphisms of $\mathbb{C}$. Of course, if $\mathbb{C}$ is already a groupoid, this is the same as $[\mathbb{C}, \mathcal{E}]$; however, if we wish to obtain $[\mathbb{C}, \mathcal{E}]$ in the case when $\mathbb{C}$ is not a groupoid, then we must form the 'lax descent topos', defined as in 3.4.11 but omitting the requirement that α be invertible. A direct proof that this is a topos may be given in the same way as for coinserters (cf. 3.4.10 above).

(*b*) Slightly less trivially, let $\mathbb{G} = (G_1 \rightrightarrows G_0)$ be a topological groupoid (that is, an internal groupoid in the cartesian category $\mathbf{Sp}$). As in Section B2.3, we may extend $\mathbb{G}$ to a simplicial object $G_\bullet$ in $\mathbf{Sp}$; and since the construction $X \mapsto \mathbf{Sh}(X)$ is functorial by A4.1.11, we obtain a simplicial object $\mathbf{Sh}(G_\bullet)$ in $\mathfrak{BTop}/\mathbf{Set}$. Thus we obtain a topos $\mathbf{Desc}(\mathbf{Sh}(G_\bullet))$, whose objects are sheaves E on G_0 equipped with a (necessarily invertible) action map $d_1^* E \to d_0^* E$ in $\mathbf{Sh}(G_1)$ satisfying the unit and cocycle conditions; given the equivalence between $\mathbf{Sh}(X)$ and $\mathbf{LH}/X$, to be established in C1.3.11, we may alternatively regard this as the

full subcategory of the internal diagram category $[\mathcal{G}, \mathbf{Sp}]$ (defined as in 2.3.11) whose objects are local homeomorphisms $E \to G_0$.

In particular, if $\mathbb{G}$ is a topological group (that is, if G_0 is the one-point space, so that sheaves on G_0 are simply (discrete) sets), then $\mathbf{Desc}(\mathbf{Sh}(G_\bullet))$ may be identified with the category $\mathbf{Cont}(G_1)$ of continuous G_1-sets which we defined (and showed to be a topos) in A2.1.6. We shall continue to use the notation $\mathbf{Cont}(\mathbb{G})$ for $\mathbf{Desc}(\mathbf{Sh}(G_\bullet))$ when $\mathbb{G}$ is a general groupoid in $\mathbf{Sp}$, and call it the topos of *continuous actions* of the groupoid $\mathbb{G}$. Note, however, that the proof given in A2.1.6 that $\mathbf{Cont}(G)$ is a topos cannot be extended to the case of a groupoid $\mathbb{G}$ whose space of objects G_0 is not discrete, since we do not have a comonadic forgetful functor from $\mathbf{Cont}(\mathbb{G})$ to the topos $[U\mathbb{G}, \mathbf{Set}]$ of 'not-necessarily-continuous actions of $\mathbb{G}$'. (Here U denotes the forgetful functor $\mathbf{Sp} \to \mathbf{Set}$.) And when we pass from topological groupoids to localic groupoids (that is, internal groupoids in the category $\mathbf{Loc}$ of locales, to be introduced in Chapter C1), we cannot even employ the argument of A2.1.6 in the case of a group. However, the argument via categories of descent data, given here, goes through without change for localic groupoids.

(c) In practice, we are often interested in studying $\mathbf{Desc}(\mathcal{E}_\bullet)$ in the case when the simplicial topos $\mathcal{E}_\bullet$ arises from the iterated pullbacks of a (bounded) morphism $f : \mathcal{E}_0 \to \mathcal{F}$; that is, we have $\mathcal{E}_1 = \mathcal{E}_0 \times_{\mathcal{F}} \mathcal{E}_0$, $\mathcal{E}_2 = \mathcal{E}_0 \times_{\mathcal{F}} \mathcal{E}_0 \times_{\mathcal{F}} \mathcal{E}_0$, and so on, with the d_i^n and s_j^n induced by the product projections and diagonal maps. In this case (as in (*b*), and the sub-case of (*a*) where $\mathbb{C}$ is a groupoid), we *do* have a twist equivalence $t : \mathcal{E}_1 \to \mathcal{E}_1$, obtained by interchanging the factors in the pullback; that is, we may extend $\mathcal{E}_\bullet$ to a pseudofunctor defined on the opposite of the category $\mathbf{Set}_f^+$ of nonempty finite sets and all maps between them. Accordingly, in defining the category $\mathbf{Desc}(\mathcal{E}_\bullet)$, we may leave out the requirement that the structure map α be invertible, since its image under t^* will be (modulo coherence isomorphisms) a two-sided inverse for it.

Suggestions for further reading: Moerdijk [831], Wraith [1235].

B4

$\mathfrak{BTop}/\mathcal{S}$ AS A 2-CATEGORY

B4.1 Finite weighted limits

In this chapter we shall study the remarkable 2-categorical structure of the 2-category $\mathfrak{BTop}/\mathcal{S}$ of bounded $\mathcal{S}$-toposes, where $\mathcal{S}$ is a fixed topos. We shall restrict ourselves entirely to bounded geometric morphisms throughout the chapter: although some of our results can be extended to unbounded morphisms, the structure of $\mathfrak{Top}/\mathcal{S}$ is much more delicate, and the gain in generality does not seem worth the extra trouble that would be involved in formulating all our results so as to apply to unbounded morphisms. In this first section, our aim is to construct finite weighted limits (in the sense of 1.1.5) in $\mathfrak{BTop}/\mathcal{S}$.

The construction of ordinary 'conical' finite limits was almost done in Section B3.3: there we saw that the pullback of a bounded geometric morphism along an arbitrary geometric morphism exists and is bounded. So, if we are given a pair of bounded morphisms $f\colon \mathcal{F} \to \mathcal{E}$ and $g\colon \mathcal{G} \to \mathcal{E}$, both projections from the pullback

$$\begin{tikzcd} \mathcal{H} \arrow[r] \arrow[d] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

are bounded; since composites of bounded morphisms are bounded by 3.1.10(i), it follows that $\mathcal{H}$ is bounded over $\mathcal{E}$, and more generally over any topos $\mathcal{S}$ such that $\mathcal{E}$ is bounded over $\mathcal{S}$. And since $\mathfrak{BTop}/\mathcal{S}$ is a full subcategory of $\mathfrak{Top}/\mathcal{S}$ by 3.1.10(ii), the square above is a pullback in it.

Since the identity morphism on $\mathcal{S}$ is trivially a terminal object for $\mathfrak{BTop}/\mathcal{S}$, we have thus established

Theorem 4.1.1 *For any topos $\mathcal{S}$, $\mathfrak{BTop}/\mathcal{S}$ has finite conical limits.* □

Thus, to complete the proof that $\mathfrak{BTop}/\mathcal{S}$ has all finite weighted limits, it suffices by 1.1.5 to prove that it has cotensors with **2**. The major part of the work is contained in

Proposition 4.1.2 *For any topos $\mathcal{S}$, the 2-functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$ sending $\mathbb{C}$ to the diagram category $[\mathbb{C}, \mathcal{S}]$ preserves cotensors with* **2**.

Proof Let us write $\mathbb{D}$ for the cotensor $\mathbf{2} \pitchfork \mathbb{C}$ in $\mathfrak{Cat}(\mathcal{S})$: of course, we have $D_0 = C_1$, and D_1 is the object $S_{\mathbb{C}}$ of commutative squares in $\mathbb{C}$, as defined in 2.6.1(*d*). Since the construction of $\mathbb{D}$ from $\mathbb{C}$ involves only finite limits, it is preserved by any inverse image functor; thus, in order to show that we have an equivalence

$$\mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{D}, \mathcal{S}]) \simeq [\mathbf{2}, \mathfrak{Top}/\mathcal{S}\,(\mathcal{E}, [\mathbb{C}, \mathcal{S}])]$$

for any $\mathcal{S}$-topos $\mathcal{E}$, it suffices by Diaconescu's Theorem to show that this equivalence holds for $\mathcal{E} = \mathcal{S}$, i.e. that we have

$$\mathbf{Tors}(\mathbb{D}, \mathcal{S}) \simeq [\mathbf{2}, \mathbf{Tors}(\mathbb{C}, \mathcal{S})]\,.$$

Now the universal 2-cell

$$\begin{tikzcd} \mathbb{D} \arrow[r, bend left, "s"] \arrow[r, bend right, "t"'] \arrow[r, phantom, "\downarrow\alpha"] & \mathbb{C} \end{tikzcd}$$

in $\mathfrak{Cat}(\mathcal{S})$ induces a 2-cell between the corresponding diagram toposes, since $\mathbb{C} \mapsto [\mathbb{C}, \mathcal{S}]$ is a 2-functor; and composition with this 2-cell yields a functor $\mathbf{Tors}(\mathbb{D}, \mathcal{S}) \to [\mathbf{2}, \mathbf{Tors}(\mathbb{C}, \mathcal{S})]$. In order to show that this functor is one half of an equivalence of categories, we must investigate its effect in more detail. Since we are now working inside a single topos $\mathcal{S}$, we shall feel free to use set-theoretic notation and terminology for its objects; the justification for this is to be found in Chapter D1.

Let $T_0, T_1 : \mathbb{C}^{\mathrm{op}} \rightrightarrows \mathcal{S}$ be two $\mathbb{C}$-torsors, and $\alpha\colon T_0 \to T_1$ a natural transformation between them. We define $U\colon \mathbb{D}^{\mathrm{op}} \to \mathcal{S}$ by the pullback diagram

$$\begin{tikzcd} U(f) \arrow[r] \arrow[d] & T_1(\mathrm{cod}\ f) \arrow[d, "T_1(f)"] \\ T_0(\mathrm{dom}\ f) \arrow[r, "\alpha_{\mathrm{dom}\ f}"] & T_1(\mathrm{dom}\ f) \end{tikzcd}$$

where f is an object of $\mathbb{D}$, i.e. a morphism of $\mathbb{C}$, with the obvious action on U of morphisms of $\mathbb{D}$. We must first show that U is a torsor, i.e. that the domain $\mathbb{U}$ of the corresponding discrete fibration $\mathbb{U} \to \mathbb{D}$ is filtered.

First, ob $\mathbb{U}$ is nonempty: since T_0 is a torsor, there exists $x \in T_0(c)$ for some $c \in \mathrm{ob}\ \mathbb{C}$, and then the pair $(x, \alpha_c(x))$ defines an element of $U(1_c)$. Next, suppose we are given elements $(x_1, y_1) \in U(f_1 : c_1 \to d_1)$,$(x_2, y_2) \in U(f_2 : c_2 \to d_2)$. Then since T_0 is a torsor there exist $g_i\colon c_i \to c_3$ $(i = 1, 2)$ and $x_3 \in T_0(c_3)$ with $T_0(g_i)(x_3) = x_i$ $(i = 1, 2)$. Let $y_3 = \alpha_{c_3}(x_3) \in T_1(c_3)$; then $T_1(g_i)(y_3) = T_1(f_i)(y_i)$ $(i = 1, 2)$, so using the fact that T_1 is a torsor we can find morphisms $h_i\colon d_i \to d_3$ $(i = 1, 2)$ and $h_3\colon c_3 \to d_3$, together with $y_4 \in T_1(d_3)$, such that $h_i f_i = h_3 g_i$ $(i = 1, 2)$ and $T_1(h_i)(y_4) = y_i$ $(i = 1, 2, 3)$. Now the pair (x_3, y_4)

defines an element of $U(h_3)$, and we have morphisms $(g_i, h_i)\colon f_i \to h_3$ with $U(g_i, h_i)(x_3, y_4) = (x_i, y_i)$ $(i = 1, 2)$. So we have verified the second condition for filteredness of $\mathbb{U}$.

For the third condition, suppose given a diagram

$$\begin{tikzcd} c_1 \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] \arrow[d, "f_1"] & c_2 \arrow[d, "f_2"] \\ d_1 \arrow[r, shift left, "h_1"] \arrow[r, shift right, "h_2"'] & d_2 \end{tikzcd}$$

and an element (x, y) of $U(f_2)$ with $T_0(g_1)(x) = T_0(g_2)(x)$ and $T_1(h_1)(y) = T_1(h_2)(y)$. Then there exists $g_3\colon c_2 \to c_3$ with $g_3 g_1 = g_3 g_2$ and an element $x' \in T_0(c_3)$ such that $T_0(g_3)(x') = x$. Now if $y' = \alpha_{c_3}(x') \in T_1(c_3)$, we have $T_1(g_3)(y') = \alpha_{c_2}(x) = T_1(f_2)(y)$, so since T_1 is a torsor we can find $h_3\colon d_2 \to d_3$ and $f_3\colon c_3 \to d_3$ with $f_3 g_3 = h_3 f_2$ and $h_3 h_1 = h_3 h_2$, and an element $y'' \in T_1(d_3)$ such that $T_1(f_3)(y'') = y'$ and $T_1(h_3)(y'') = y$. Thus the pair (x', y'') defines an element of $U(f_3)$ satisfying $U(g_3, h_3)(x', y'') = (x, y)$. So we have verified that U is a torsor.

Next, we must verify that if we apply the functor $\mathbf{Tors}(\mathbb{D}, \mathcal{S}) \to [\mathbf{2}, \mathbf{Tors}(\mathbb{C}, \mathcal{S})]$ which we defined earlier to U, we recover the transformation $\alpha\colon T_0 \to T_1$ up to isomorphism. In particular, we must verify the coend formulae

$$T_0(c) \cong \int^f \mathbb{C}(c, \mathrm{dom}\ f) \times U(f)$$

and

$$T_1(c) \cong \int^f \mathbb{C}(c, \mathrm{cod}\ f) \times U(f)\,.$$

The first of these is easy, because the coequalizer diagram defining the coend splits to yield an isomorphism $\int^f \mathbb{C}(c, \mathrm{dom}\ f) \times U(f) \cong U(1_c)$, and it is clear from the definition of U that we have $U(1_c) \cong T_0(c)$. For the second, note that we have a dinatural map

$$\coprod_f \mathbb{C}(c, \mathrm{cod}\ f) \times U(f) \longrightarrow T_1(c)$$

sending $(g, (x, y))$ to $T_1(g)(y)$, and hence an induced map ϕ from the coend to $T_1(c)$. This map is surjective: given $z \in T_1(c)$, there exists $x \in T_0(d)$ for some d, and then since T_1 is a torsor we can find $f\colon d \to e$, $g\colon c \to e$ and $y \in T_1(e)$ such that $T_1(g)(y) = z$ and $T_1(f)(y) = \alpha_d(x)$; so the element of the coend represented by $(g, (x, y))$ is mapped by ϕ to z.

To show that ϕ is injective, suppose given elements $(x_i, y_i) \in U(f_i : d_i \to e_i)$ $(i = 1, 2)$ and morphisms $g_i : c \to e_i$ such that $T_1(g_1)(y_1) = T_1(g_2)(y_2)$. We have to show that $(g_1, (x_1, y_1))$ and $(g_2, (x_2, y_2))$ represent the same element of the coend. But we can find $h_i : d_i \to d_3$ $(i = 1, 2)$ and $x_3 \in T_0(d_3)$ with $T_0(h_i)(x_3) = x_i$ $(i = 1, 2)$; and then for $y_3 = \alpha_{d_3}(x_3)$ we have $T_1(h_i)(y_3) = \alpha_{d_i}(x_i) = T_1(f_i)(y_i)$ $(i = 1, 2)$, as well as $T_1(g_1)(y_1) = T_1(g_2)(y_2)$. So we can find $k_i : e_i \to e_3$ $(i = 1, 2)$ and $k_3 : d_3 \to e_3$ with $k_3 h_i = k_i f_i$ $(i = 1, 2)$ and $k_1 g_1 = k_2 g_2$, together with an element $y_4 \in T_1(e_3)$ with $T_1(k_i)(y_4) = y_i$ $(i = 1, 2)$ and $T_1(k_3)(y_4) = y_3$. Then it is clear that the element $(k_1 g_1, (x_3, y_4))$ of $\mathbb{C}(c, \text{cod } k_3) \times U(k_3)$ represents the same element of the coend as both $(g_1, (x_1, y_1))$ and $(g_2, (x_2, y_2))$. So we have identified the domain and codomain of the natural transformation

$$\int^f \mathbb{C}(c, \text{dom } f) \times U(f) \longrightarrow \int^f \mathbb{C}(c, \text{cod } f) \times U(f)$$

(which sends the element represented by $(g, (x, y))$ to that represented by $(fg, (x, y)))$ with $T_0(c)$ and $T_1(c)$ respectively; it is now straightforward to identify the morphism itself with α_c.

Finally, we must verify that if we start with an arbitrary $\mathbb{D}$-torsor U and construct $\alpha : T_0 \to T_1$ by the coend formulae above, then we may reconstruct U up to isomorphism by the pullback square

$$\begin{tikzcd}
U(f) \arrow[r] \arrow[d] & T_1(\text{cod } f) \arrow[d, "T_1(f)"] \\
T_0(\text{dom } f) \arrow[r, "\alpha_{\text{dom } f}"] & T_1(\text{dom } f)
\end{tikzcd}$$

But we certainly have a commutative square as above: the top edge sends $z \in U(f)$ to the element of the coend represented by $(1_{\text{cod } f}, z)$, and the left-hand edge sends it to the element represented by $(1_{\text{dom } f}, z)$. So we have to show that the induced map from $U(f)$ to the pullback is bijective.

For surjectivity, suppose $f : c \to d$ and we are given morphisms

$$d_2 \xleftarrow{h_2} c_2 \xleftarrow{g_2} c \xrightarrow{f} d \xrightarrow{g_1} d_1 \xleftarrow{h_1} c_1$$

and elements $z_i \in U(h_i)$ $(i = 1, 2)$, so that the pairs (g_1, z_1) and (g_2, z_2) represent elements of $T_1(d)$ and $T_0(c)$ respectively. Then, since U is a torsor, the compatibility condition $T_1(f)(g_1, z_1) = \alpha_c(g_2, z_2)$ is equivalent to saying that we

have a commutative diagram

$$\begin{tikzcd}
c_2 \arrow[r, "k_2"] \arrow[d, "h_2"] & c_3 \arrow[d, "h_3"] & c_1 \arrow[l, "k_1"'] \arrow[d, "h_1"] \\
d_2 \arrow[r, "l_2"] & d_3 & d_1 \arrow[l, "l_1"'] \\
& c \arrow[ul, "h_2 g_2"] \arrow[ur, "g_1 f"'] &
\end{tikzcd}$$

and an element $z_3 \in U(h_3)$ such that $U(k_i, l_i)(z_3) = z_i$ $(i = 1, 2)$. But now $(k_2 g_2, l_1 g_1)$ is a morphism $f \to h_3$ in $\mathbb{D}$, and the element $U(k_2 g_2, l_1 g_1)(z_3)$ of $U(f)$ is easily seen to map to the two given elements of $T_1(d)$ and $T_0(c)$.

For injectivity, suppose we are given two elements z_1, z_2 of $U(f : c \to d)$ such that $(1_d, z_1)$ and $(1_d, z_2)$ represent the same element of $T_1(d)$, and $(1_c, z_1)$ and $(1_c, z_2)$ represent the same element of $T_0(c)$. By the identification of $T_0(c)$ with $U(1_c)$, the latter condition is equivalent to saying that $U(1_c, f)(z_1) = U(1_c, f)(z_2)$ in $U(1_c)$; since U is a torsor, this implies that there are morphisms $(g_i, h_i) : f \to f'$ $(i = 1, 2)$ in $\mathbb{D}$ such that $g_1 = g_2$ (and $h_1 f = h_2 f$), and an element $z_3 \in U(f')$ such that $U(g_i, h_i)(z_3) = z_i$ $(i = 1, 2)$. But now the pair (h_i, z_3) represents the same element of $T_1(d)$ as $(1_d, z_i)$ $(i = 1, 2)$, so from the fact that these two are equal we deduce that there exists $(g_3, h_3) : f' \to f''$ and $z_4 \in U(f'')$ such that $h_3 h_1 = h_3 h_2$ and $U(g_3, h_3)(z_4) = z_3$. So $z_1 = U(g_3 g_1, h_3 h_1)(z_4) = U(g_3 g_2, h_3 h_2)(z_4) = z_2$. □

Remark 4.1.3 As well as cotensors with $\mathbf{2}$, we know that the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{BTop}/\mathcal{S}$ of 4.1.2 preserves finite products, by 3.2.13. However, it does not preserve all finite limits. The following counterexample borrows ideas from the theory of classifying toposes, and so properly belongs in the next section; but there seems to be little harm in giving it here.

Let $\mathbf{Set}_f$ denote (a small skeleton of) the category of finite sets, and consider the equalizer of

$$[\mathbf{2}, \mathbf{Set}_f] \underset{d_1}{\overset{d_0}{\rightrightarrows}} \mathbf{Set}_f$$

in $\mathfrak{Cat}(\mathbf{Set})$. By definition (remember that we are dealing with pseudo-limits!), the objects of this category consist of objects $(f : A_0 \to A_1)$ of $[\mathbf{2}, \mathbf{Set}_f]$ equipped with an additional isomorphism $g : A_0 \to A_1$; but this object is (canonically) isomorphic to the object $(g^{-1} f : A_0 \to A_0)$ equipped with the identity isomorphism, and hence the equalizer is equivalent to the category of finite sets equipped with

an endomorphism. (Warning: it is not true in general that strict equalizers and pseudo-equalizers in $\mathfrak{Cat}(\mathbf{Set})$ yield equivalent categories.) The functor category $[\mathbf{Set}_f, \mathbf{Set}]$ is an object classifier over **Set**, by 3.2.9: that is, for any **Set**-topos $\mathcal{E}$ we have

$$\mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, [\mathbf{Set}_f, \mathbf{Set}]) \simeq \mathcal{E}\,.$$

Hence by 4.1.2 $[[\mathbf{2}, \mathbf{Set}_f], \mathbf{Set}] \simeq \mathbf{2} \pitchfork [\mathbf{Set}_f, \mathbf{Set}]$ is a morphism classifier. Taking the equalizer in $\mathfrak{BTop}/\mathbf{Set}$ of the geometric morphisms induced by d_0 and d_1, we therefore obtain (by an argument like that already given) a classifying topos for the theory whose models in $\mathcal{E}$ are objects of $\mathcal{E}$ equipped with an endomorphism. But this is a finitary algebraic theory, and the classifying topos for such a theory is well known to be of the form $[\mathcal{C}, \mathbf{Set}]$ where $\mathcal{C}$ is the category of finitely presented models of the theory in **Set** (cf. D3.1.2). Since not every finitely presented model of the theory is finite (in particular, the free model on one generator may be identified with the set of natural numbers with the successor map), this topos is not equivalent to $[\mathcal{C}_0, \mathbf{Set}]$ where $\mathcal{C}_0$ is the equalizer computed in $\mathfrak{Cat}(\mathbf{Set})$. (In fact the latter can be shown to be a classifying topos for 'torsion models' of the theory, that is those in which every element generates a finite submodel.)

Despite the foregoing, the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{BTop}/\mathcal{S}$ does preserve arbitrary comma objects $(f \downarrow g)$. This may be proved by a straightforward generalization of the argument in the proof of 4.1.2 (the latter being, of course, the special case in which f and g are identity functors).

Theorem 4.1.4 *For any topos $\mathcal{S}$, $\mathfrak{BTop}/\mathcal{S}$ has cotensors with* **2**.

Proof Let $\mathcal{E}$ be a bounded $\mathcal{S}$-topos. If $\mathcal{E}$ is of the form $[\mathbb{C}, \mathcal{S}]$, the existence of the cotensor $\mathbf{2} \pitchfork \mathcal{E}$ follows immediately from 4.1.2. In general, we have an inclusion $i\colon \mathcal{E} \to \mathcal{F}$ where $\mathcal{F}$ is of the above form; then for any topos $\mathcal{G}$ the functor

$$\mathfrak{Top}\,(\mathcal{G}, \mathcal{E}) \longrightarrow \mathfrak{Top}\,(\mathcal{G}, \mathcal{F})$$

induced by composition with i is a full embedding, as we noted before 1.1.19. And this clearly remains true if we replace the 2-category $\mathfrak{Top}$ by $\mathfrak{BTop}/\mathcal{S}$. So, if we form the pullback

$$\begin{tikzcd} \mathcal{P} \arrow[r] \arrow[d] & \mathcal{E} \times \mathcal{E} \arrow[d, "i \times i"] \\ \mathbf{2} \pitchfork \mathcal{F} \arrow[r, "(d_0, d_1)"] & \mathcal{F} \times \mathcal{F} \end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$, it is easily verified that $\mathcal{P}$ has the universal property of a cotensor $\mathbf{2} \pitchfork \mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$. □

Remark 4.1.5 As we observed in Section B3.4, we may regard $\mathfrak{BTop}/\mathcal{S}$ as a $\mathfrak{Cat}_{\mathcal{S}}$-indexed category; so it makes sense to talk about cotensors with internal categories in $\mathcal{S}$, not just with finite 'external' categories such as $\mathbf{2}$. And in fact we may show that all such cotensors exist (at least provided $\mathcal{S}$ has a natural number object), using the theory of classifying toposes to be developed in the next section: if $\mathcal{E}$ is a bounded $\mathcal{S}$-topos, classifying a geometric theory $\mathbb{T}$ over $\mathcal{S}$, then $\mathbb{C} \pitchfork \mathcal{E}$ may be defined as the classifying topos for the theory of diagrams of $\mathbb{T}$-models of shape $\mathbb{C}$, and a presentation of the latter as a geometric theory may easily be constructed from one for $\mathbb{T}$. On the other hand, the functor $\mathfrak{Cat}(\mathcal{S}) \rightarrow \mathfrak{BTop}/\mathcal{S}$ does not in general preserve these more general cotensors (that is, $\mathbb{C} \pitchfork [\mathbb{D}, \mathcal{S}]$ is not equivalent to $[\mathbb{D}^{\mathbb{C}}, \mathcal{S}]$): the reason is that, although the hard part of the proof of 4.1.2 still works, the construction of the exponential $\mathbb{D}^{\mathbb{C}}$ in $\mathbf{Cat}(\mathcal{S})$ (cf. 2.3.15(ii)) involves exponentials in $\mathcal{S}$, which are not preserved by inverse image functors.

We shall also see later in this chapter that $\mathfrak{BTop}/\mathcal{S}$ has $\mathcal{S}$-indexed products in an apprpopriate sense, and not just finite ones; cf. 4.3.4 below. Hence it also has arbitrary $\mathcal{S}$-indexed conical limits, since these can be constructed from products and equalizers. In Part C, we shall prove a number of results about cofiltered $\mathcal{S}$-indexed limits, that is limits indexed by cofiltered internal categories in $\mathcal{S}$.

The proofs of 4.1.1 and 4.1.4 both made use of the explicit representation of bounded $\mathcal{S}$-toposes provided by Giraud's theorem. However, there are some important classes of weighted limits, in particular inverters and equifiers (cf. 1.1.4(e) and (f)), which can be constructed directly from the definition of a bounded morphism, using the machinery of A4.5.13.

Lemma 4.1.6 *Let $p\colon \mathcal{E} \rightarrow \mathcal{S}$ be a bounded $\mathcal{S}$-topos. Then there is an object E of $\mathcal{E}$ such that, for any 2-cell*

$$\begin{tikzcd} \mathcal{F} \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}/\mathcal{S}$ with codomain $\mathcal{E}$, α is an isomorphism iff its component α_E is an isomorphism.

Proof Let B be a bound for $\mathcal{E}$ over $\mathcal{S}$. Form the pullback

$$\begin{tikzcd} E \arrow[r] \arrow[d, tail] & \in_B \arrow[d, tail] \\ p^*p_*(PB) \times B \arrow[r, "\epsilon_{PB} \times 1"] & PB \times B \end{tikzcd}$$

where PB is the power object of B, and ϵ is the counit of $(p^* \dashv p_*)$. We shall show that if α_E is an isomorphism then α_A is an isomorphism for all objects A of $\mathcal{E}$.

First consider an object A of $\mathcal{E}$ which is a subobject of $p^*I \times B$ for some object I of $\mathcal{S}$. Let $h\colon I \to p_*(PB)$ be the transpose across $(p^* \dashv p_*)$ of the name of the relation $A \rightarrowtail p^*I \times B$; then there is a pullback diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d, tail] & E \arrow[d, tail] \\ p^*I \times B \arrow[r, "p^*(h) \times 1"] & p^*p_*(PB) \times B \end{tikzcd}$$

and hence a pullback diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d] & E \arrow[d] \\ p^*I \arrow[r, "p^*(h)"] & p^*p_*(PB) \end{tikzcd} .$$

Now α_E is an isomorphism by assumption; and α_{p^*I} is the canonical isomorphism $f^*p^*I \cong q^*I \cong g^*p^*I$, since α is a geometric transformation over $\mathcal{S}$, and similarly for $\alpha_{p^*p_*(PB)}$. So the components of α at three corners of the pullback square above are isomorphisms; and since the pullback square is preserved by both f^* and g^*, it follows that α_A must be an isomorphism.

Now suppose A is an arbitrary object of $\mathcal{E}$. By the definition of a bounded morphism, we can find an epimorphism $C \twoheadrightarrow A$, where C is a subobject of $p^*I \times B$ for some I. Let $R \rightrightarrows C$ be the kernel-pair of this epimorphism, and choose an epimorphism $D \twoheadrightarrow R$ where D is a subobject of some $p^*J \times B$. Then we have a coequalizer diagram $D \rightrightarrows C \twoheadrightarrow A$, and we know that α_C and α_D are isomorphisms by the argument above. But, once again, this coequalizer is preserved by both f^* and g^*; so it follows that α_A must be an isomorphism. □

Corollary 4.1.7 *The* 2*-category* $\mathfrak{BTop}/\mathcal{S}$ *has inverters and equifiers.*

Proof To form the inverter of a geometric transformation

$$\begin{tikzcd} \mathcal{F} \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & \mathcal{E}, \end{tikzcd}$$

let E be an object of $\mathcal{E}$ as in 4.1.6, and let j be the smallest local operator on $\mathcal{F}$ for which $L(\alpha_E)$ is an isomorphism (where L is the associated j-sheaf

functor), as constructed in A4.5.14(c). Then a geometric morphism $h\colon \mathcal{G} \to \mathcal{F}$ factors through $\mathbf{sh}_j(\mathcal{F}) \to \mathcal{F}$ iff the local operator k corresponding to its image satisfies $k \geq j$, iff $h^*(\alpha_E)$ is an isomorphism. So it follows from 4.1.6 that the inclusion $\mathbf{sh}_j(\mathcal{F}) \to \mathcal{F}$ is the required inverter.

To form the equifier of a parallel pair of geometric transformations $\alpha, \beta\colon f \rightrightarrows g$, we similarly construct the smallest local operator j on $\mathcal{F}$ for which the associated sheaf functor L maps the equalizer of α_E and β_E to an isomorphism. Then $L(\alpha_E) = L(\beta_E)$; and by an argument just like that in the proof of 4.1.6 it follows that $L(\alpha_A) = L(\beta_A)$ for every object A of $\mathcal{F}$, and hence that $\mathbf{sh}_j(\mathcal{F}) \to \mathcal{F}$ is an equifier for α and β. □

We shall make extensive use of inverters in $\mathfrak{BTop}/\mathcal{S}$ in Section B4.5 below. Also in that section, we shall require a remarkable result about comma objects in $\mathfrak{BTop}/\mathcal{S}$, originally due to A. M. Pitts. What is surprising about it is its 'something for nothing' aspect: starting from a mere $\mathcal{S}$-indexed left adjoint, we end by constructing one which is not only $\mathcal{S}$-indexed but $\mathcal{G}$-indexed, where $\mathcal{G}$ is a topos over $\mathcal{S}$. The proof of this result involves working with sites, so we defer it to Part C (specifically, to C2.3.17); but we state the result here, so that we can refer to it in Section B4.5. Recall from 3.1.1 that a geometric morphism over $\mathcal{S}$ is said to be $\mathcal{S}$-essential if its inverse image functor has an $\mathcal{S}$-indexed left adjoint.

Theorem 4.1.8 *Let*

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\to"] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

be a comma square in $\mathfrak{BTop}/\mathcal{S}$, and suppose f is $\mathcal{S}$-essential. Then q is $\mathcal{G}$-essential, and the Beck–Chevalley natural transformation $q_!p^* \to g^*f_!$ (the mate of the (non-invertible!) natural transformation $p^*f^* \to q^*g^*$) is an isomorphism. □

There is one particular case in which we can prove 4.1.8 with the tools currently at our disposal: namely that in which f and g are induced by functors between internal categories in $\mathcal{S}$. For we noted in 4.1.3 that the comma square

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\to"] & \mathbb{D} \arrow[d, "g"] \\
\mathbb{C} \arrow[r, "f"] & \mathbb{B}
\end{tikzcd}$$

in $\mathfrak{Cat}(\mathcal{S})$ is preserved by the functor $\mathfrak{Cat}(\mathcal{S}) \to \mathfrak{BTop}/\mathcal{S}$ of 2.3.22; and the top edge of this square is a split opfibration, i.e. it corresponds to a (strict) functor $\mathbb{D} \to \mathbf{Cat}(\mathcal{S})$. (In set-theoretic terms, it is the functor which sends an object D to the comma category $(f \downarrow D)$, and a morphism $\delta\colon D \to D'$ to the functor induced by composition with δ.) Thus we can identify the morphism $[(f \downarrow g), \mathcal{S}] \to [\mathbb{D}, \mathcal{S}]$ induced by q with one of the form $[\mathbb{A}, [\mathbb{D}, \mathcal{S}]] \to [\mathbb{D}, \mathcal{S}]$ where $\mathbb{A}$ is an internal category in $[\mathbb{D}, \mathcal{S}]$; and the latter is clearly $[\mathbb{D}, \mathcal{S}]$-essential (just as $[\mathbb{C}, \mathcal{S}] \to [\mathbb{B}, \mathcal{S}]$ is automatically $\mathcal{S}$-essential in this case), by the remark after 3.1.1. Also, we noted in 2.5.11 that the Beck–Chevalley condition holds for comma squares of this type.

B4.2 Classifying toposes via weighted limits

One of the central themes of topos theory is the construction of classifying toposes for geometric theories, and the study of their properties. Since this properly belongs to the interaction between toposes and (first-order) logic, it will be treated in detail in Part D – specifically, in Chapter D3. However, there is an alternative presentation of the construction, which makes it a part of the 2-categorical study of $\mathfrak{Top}$, and which may appropriately be described at this point. We shall not presume any familiarity with the formal definition of a geometric theory (which can be found in Chapter D1); for this presentation, it will suffice to take a more informal notion of a 'theory' as a functor on toposes over a given base.

Definition 4.2.1 (*a*) Let $\mathcal{S}$ be a topos. By a *theory* $\mathbb{T}$ over $\mathcal{S}$, we mean (informally) a $(\mathfrak{BTop}/\mathcal{S})$-indexed category, i.e. a (pseudo-)functor $\mathbb{T}\colon (\mathfrak{BTop}/\mathcal{S})^{\mathrm{op}} \to \mathfrak{CAT}$. That is, $\mathbb{T}$ assigns to each bounded $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ a category $\mathbb{T}(\mathcal{E})$ (whose objects are called $\mathbb{T}$-*models* in $\mathcal{E}$), to each geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ a functor $f^*\colon \mathbb{T}(\mathcal{E}) \to \mathbb{T}(\mathcal{F})$, and to each geometric transformation $\alpha\colon f \to g$ over $\mathcal{S}$ a natural transformation $\alpha\colon f^* \to g^*$, subject to the obvious compatibility conditions.

(*b*) By a *classifying topos* for a theory $\mathbb{T}$ over $\mathcal{S}$, we mean a representing object (in the up-to-equivalence sense) for the functor $\mathbb{T}$; that is, a bounded $\mathcal{S}$-topos $\mathcal{S}[\mathbb{T}]$ equipped with a distinguished object G of $\mathbb{T}(\mathcal{S}[\mathbb{T}])$ (the *generic* $\mathbb{T}$-*model over* $\mathcal{S}$) such that, for each bounded $\mathcal{S}$-topos $\mathcal{E}$, the functor

$$\mathfrak{BTop}/\mathcal{S}\,(\mathcal{E}, \mathcal{S}[\mathbb{T}]) \longrightarrow \mathbb{T}(\mathcal{E})$$

which sends a geometric morphism f to $f^*(G)$ (and a geometric transformation α to α_G), is one half of an equivalence of categories.

Clearly the classifying topos of a theory $\mathbb{T}$, if it exists, is unique up to canonical equivalence in $\mathfrak{BTop}/\mathcal{S}$. We note that the existence of a classifying topos for $\mathbb{T}$ imposes some restrictions on $\mathbb{T}$: in particular, it must transform arbitrary

(weighted) colimits in $\mathfrak{BTop}/\mathcal{S}$ to limits in $\mathfrak{CAT}$. Applying this to the results of 3.4.1 and 1.5.6, we obtain

Lemma 4.2.2 *If $\mathbb{T}$ has a classifying topos over $\mathcal{S}$, then for any $\mathcal{S}$-topos $\mathcal{E}$ the $\mathcal{E}$-indexed category $\mathbb{T}_{\mathcal{E}}$ obtained by restricting $\mathbb{T}$ to toposes of the form $\mathcal{E}/A$ (and geometric morphisms over $\mathcal{E}$ between them, cf. 3.2.8(b)) is a stack for the coherent coverage on $\mathcal{E}$.*

Proof For any two objects A and B of $\mathcal{E}$, we have $\mathcal{E}/A \amalg B \simeq \mathcal{E}/A \times \mathcal{E}/B$, since coproducts in $\mathcal{E}$ are stable under pullback; so 3.4.1 implies that $\mathbb{T}(\mathcal{E}/A \amalg B) \simeq \mathbb{T}(\mathcal{E}/A) \times \mathbb{T}(\mathcal{E}/B)$. Similarly, if $f: A \twoheadrightarrow B$ is an epimorphism in $\mathcal{E}$ with kernel-pair $a, b: R \rightrightarrows A$, then 1.5.6(iii) implies that we may identify $\mathcal{E}/B$ with the category $\mathbf{Desc}(\mathbb{E}, (f))$ of objects $g: X \to A$ of $\mathcal{E}/A$ equipped with descent data $\alpha: a^*X \to b^*X$ relative to f. So, if a geometric morphism $h: \mathcal{E}/A \to \mathcal{F}$ is 'equipped with descent data for f', in the sense that h^* lifts to a functor $\mathcal{F} \to \mathbf{Desc}(\mathbb{E}, (f))$, then it factors uniquely through $\mathcal{E}/A \to \mathcal{E}/B$. Taking $\mathcal{F}$ to be the classifying topos for $\mathbb{T}$, this says that there is an equivalence $\mathbb{T}(\mathcal{E}/B) \simeq \mathbf{Desc}(\mathbb{T}_{\mathcal{E}}, (f))$. □

Similarly, we have

Lemma 4.2.3 *If $\mathbb{T}$ has a classifying topos over $\mathcal{S}$, then for any $\mathcal{S}$-topos $\mathcal{E}$ we have $\mathbb{T}([\mathbf{2}, \mathcal{E}]) \simeq [\mathbf{2}, \mathbb{T}(\mathcal{E})]$.*

Proof In 3.4.2 we saw that $[\mathbf{2}, \mathcal{E}]$ is the tensor $\mathbf{2} \otimes \mathcal{E}$ in $\mathfrak{Top}$ (and hence in $\mathfrak{Top}/\mathcal{S}$, if $\mathcal{E}$ is defined over $\mathcal{S}$); so $\mathbb{T}$ must map it to a cotensor in $\mathfrak{CAT}$. □

Lemma 4.2.3 implies that, if we know what the objects of $\mathbb{T}(\mathcal{E})$ are for all bounded $\mathcal{S}$-toposes $\mathcal{E}$ (and $\mathbb{T}$ is 'classifiable'), then we must know the morphisms as well. For example, from the fact that the theory of objects is classifiable (cf. 4.2.4(a) below), we can deduce that there does *not* exist a classifying topos for the theory defined by taking $\mathbb{T}(\mathcal{E})$ to be the category of all objects of $\mathcal{E}$ and monomorphisms between them (even though this theory is easily seen to satisfy the criterion of 4.2.2).

Before embarking on the general construction of classifying toposes, it seems appropriate to give some examples of the theories we have in mind.

Examples 4.2.4 (*a*) The theory $\mathbb{O}$ of *objects* is simply the forgetful functor which sends an $\mathcal{S}$-topos to its underlying category, a geometric morphism to its inverse image functor and a geometric transformation to itself. We saw in 3.2.9 above that this theory has a classifying topos over any topos $\mathcal{S}$ with a natural number object, namely the diagram category $[\mathbb{S}_f, \mathcal{S}]$.

(*b*) More generally, if $\mathcal{D}$ is a category, the theory $\mathbb{D}$ of *diagrams of shape* $\mathcal{D}$ sends $\mathcal{E}$ to the functor category $[\mathcal{D}, \mathcal{E}]$, and a geometric morphism f to the functor obtained from composition with f^*. In the particular case when $\mathcal{D}$ is the

finite category

$$\bullet \longrightarrow \bullet,$$

we shall denote this theory by $\mathbb{M}$ and call it the *theory of morphisms*. And in the particular case when $\mathcal{D}$ is the finite discrete category with m objects, we shall denote it by $\mathbb{O}_m$ and call it the *theory of m-tuples of objects.*

(c) As a further generalization of (*b*), if $\mathbb{T}$ is any theory over $\mathcal{S}$ and $\mathcal{D}$ is a category, we may construct a theory $[\mathcal{D}, \mathbb{T}]$ whose models are diagrams of shape $\mathcal{D}$ in the category of $\mathbb{T}$-models.

(*d*) Let I be an object of $\mathcal{S}$. The theory $\mathbb{O}_I$ of *I-indexed families of objects* is given by $\mathbb{O}_I(\mathcal{E}) = \mathcal{E}/I$ (by which, as usual, we mean $\mathcal{E}/p^*I$, where $p: \mathcal{E} \to \mathcal{S}$ is the structure morphism of $\mathcal{E}$), and $f^* : \mathbb{O}_I(\mathcal{E}) \to \mathbb{O}_I(\mathcal{F})$ is f^* applied to objects over p^*I. Clearly, the theory $\mathbb{O}_m$ defined in (*b*) above may be identified with the special case of $\mathbb{O}_I$ where I is an m-fold copower of 1. Similarly, we may form the theory $\mathbb{M}_I$ of I-indexed families of morphisms, and so on.

(*e*) Again, let I be an object of the base topos $\mathcal{S}$. Then we may construct a theory $\mathbb{S}\mathsf{ub}_I$ whose models in $\mathcal{E}$ are subobjects of p^*I (with inclusions between subobjects as morphisms). We saw in 3.2.10 above that this theory has a classifying topos over any topos $\mathcal{S}$, namely the diagram category $[\mathbb{K}(I), \mathcal{S}]$.

(*f*) Another generalization of (*b*) above: if $\mathbb{C}$ is an internal category in $\mathcal{S}$, then we have a theory whose models in $\mathcal{E}$ are diagrams of shape $\mathbb{C}$ (that is, of shape $p^*(\mathbb{C})$) in $\mathcal{E}$. Similarly, we have a theory $\mathbf{Tors}(\mathbb{C}, -)$ whose models are $\mathbb{C}$-torsors in $\mathcal{S}$-toposes. Of course, Diaconescu's Theorem 3.2.7 is exactly the statement that the latter theory has a classifying topos, namely $[\mathbb{C}, \mathcal{S}]$. And 3.2.6 combined with 3.2.7 shows that the theory of diagrams of shape $\mathbb{C}$ has a classifying topos provided $\mathcal{S}$ has a natural number object, namely the topos $[(\mathrm{Cart}(\mathbb{C}))^{\mathrm{op}}, \mathcal{S}]$ where $\mathrm{Cart}(\mathbb{C})$ is the cartesian category freely generated by $\mathbb{C}$.

In order to build more complicated examples of theories, we shall make use of the notion of a geometric construct in a theory, which we now define.

Definition 4.2.5 Let $\mathbb{T}$ be a theory over $\mathcal{S}$, as defined in 4.2.1(*a*). By a *geometric construct* in $\mathbb{T}$, we mean a $(\mathfrak{BTop}/\mathcal{S})$-indexed functor $F: \mathbb{T} \to \mathbb{O}$, where $\mathbb{O}$ is the theory of objects defined in 4.2.4(*a*). Informally, F may be thought of as something which 'constructs' an object $F(M)$ of $\mathcal{E}$ from a $\mathbb{T}$-model M in an $\mathcal{S}$-topos $\mathcal{E}$, and a morphism $F(M) \to F(N)$ from each morphism of $\mathbb{T}$-models $M \to N$, in a manner which is compatible with composition of $\mathbb{T}$-model morphisms and with the functors f^* between $\mathbb{T}$-models in different $\mathcal{S}$-toposes. A *morphism of geometric constructs* $\alpha: F \to G$ is a $(\mathfrak{BTop}/\mathcal{S})$-indexed natural transformation; equivalently, it may be thought of as a $(\mathfrak{BTop}/\mathcal{S})$-indexed functor from $\mathbb{T}$ to the theory $\mathbb{M}$ of 4.2.4(*b*).

If $\mathbb{T}$ has a classifying topos over $\mathcal{S}$, then a geometric construct F in $\mathbb{T}$ is determined up to isomorphism by the object $F(G)$ of $\mathcal{S}[\mathbb{T}]$, where G is the generic $\mathbb{T}$-model; for, given any $\mathbb{T}$-model M is an $\mathcal{S}$-topos $\mathcal{E}$, we have $F(M) \cong f^*(F(G))$, where f is the geometric morphism such that $f^*(G) \cong M$. In fact this defines an

equivalence between $\mathcal{S}[\mathbb{T}]$ and the category of geometric constructs in $\mathbb{T}$, by the '2-dimensional Yoneda lemma'. Provided the theory $\mathbb{O}$ also has a classifying topos, we may also identify geometric constructs in $\mathbb{T}$ with geometric morphisms $\mathcal{S}[\mathbb{T}] \to \mathcal{S}[\mathbb{O}]$ over $\mathcal{S}$. However, a more 'concrete' feel for what geometric constructs look like may be provided by the following examples:

Examples 4.2.6 (*a*) Let I be an object of the base topos $\mathcal{S}$. The 'constant construct' which assigns to every $\mathbb{T}$-model in an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ the object p^*I is clearly geometric. In particular, if $\mathcal{S}$ has a natural number object, then the constant functor sending a $\mathbb{T}$-model to the natural number object (in the topos where it lives) is a geometric construct.

(*b*) If n is a natural number, then the functor $A \mapsto A^n$ is a geometric construct in the theory $\mathbb{O}$ of objects, since inverse image functors preserve nth powers. More generally, if $F_1, F_2, \ldots, F_n$ are all geometric constructs in a theory $\mathbb{T}$, then so is their (pointwise) product; so, for the theory $\mathbb{D}$ of 4.2.4(*b*), we have a geometric construct sending a diagram $G\colon \mathcal{D} \to \mathcal{E}$ to $\prod_{i=1}^{n} G(f(i))$, for any function $f\colon \{1, 2, \ldots, n\} \to \mathrm{ob}\,\mathcal{D}$.

(c) Still more complicated finite limits and colimits may also be obtained as geometric constructs. For example, let G be a group in a topos $\mathcal{E}$. For any natural number n, the object G_n of elements of G of order dividing n may be constructed as the equalizer of the nth power map $G \to G$ and the constant map $G \to 1 \xrightarrow{e} G$. Clearly, this equalizer is preserved by any cartesian functor (in particular, by any inverse image functor); and it is functorial with respect to group homomorphisms in $\mathcal{E}$. So we can regard it as a geometric construct in the theory of groups, over any base topos $\mathcal{S}$. If $\mathcal{S}$ (and hence $\mathcal{E}$) has a natural number object, then we may form 'the union of the G_n for all n' as a subobject of $G \times N$; and by taking the image of this subobject under the first projection we obtain the subobject $G_t \rightarrowtail G$ of torsion elements of G; once again, this is functorial with respect to group homomorphisms in $\mathcal{E}$, and preserved by inverse image functors, and so it is a geometric construct.

Definition 4.2.7 (*a*) Let $\mathbb{T}$ be a theory over $\mathcal{S}$. A theory $\mathbb{T}'$ is called a *simple functional extension* of $\mathbb{T}$ if there exist geometric constructs F and F' in $\mathbb{T}$ such that, for every $\mathcal{S}$-topos $\mathcal{E}$, $\mathbb{T}'(\mathcal{E})$ is the category whose objects are pairs (M, f) where M is a $\mathbb{T}$-model in $\mathcal{E}$ and $f\colon FM \to F'M$ in $\mathcal{E}$, and whose morphisms are $\mathbb{T}$-model homomorphisms $h\colon M \to M'$ such that

$$\begin{tikzcd} FM \arrow[r, "f"] \arrow[d, "Fh"] & F'M \arrow[d, "F'h"] \\ FM' \arrow[r, "f'"] & F'M' \end{tikzcd}$$

commutes; and the effect of $\mathbb{T}'$ on geometric morphisms and geometric transformations is induced by that of $\mathbb{T}$ in the obvious way.

(*b*) Let $\mathbb{T}$ be a theory over $\mathcal{S}$. A theory $\mathbb{T}'$ is called a *simple geometric quotient* of $\mathbb{T}$ if there is a morphism $f: F \to F'$ of geometric constructs in $\mathbb{T}$ such that, for every $\mathcal{S}$-topos $\mathcal{E}$, $\mathbb{T}'(\mathcal{E})$ is the full subcategory of $\mathbb{T}(\mathcal{E})$ on those $\mathbb{T}$-models M for which $f_M: FM \to F'M$ is an isomorphism, and the effect of $\mathbb{T}'$ on geometric morphisms and geometric transformations is induced by that of $\mathbb{T}$ in the obvious way. (We sometimes refer to the assertion that f_M is invertible, for some particular $\mathbb{T}$-model M, as a *geometric axiom* satisfied by M.)

(*c*) We say $\mathbb{T}$ is a *geometric theory over* $\mathcal{S}$ if there is a finite sequence of theories $(\mathbb{T}_0, \mathbb{T}_1, \ldots, \mathbb{T}_n)$ such that $\mathbb{T}_0$ is the theory of m-tuples of objects for some finite m, $\mathbb{T}_n$ is equivalent to $\mathbb{T}$, and each $\mathbb{T}_{i+1}$ is either a simple functional extension or a simple geometric quotient of $\mathbb{T}_i$.

In D1.1.6 we give a very different-looking definition of the notion of geometric theory; but it turns out to have exactly the same expressive power as the one given above (that is, given a geometric theory in the above sense over **Set**, there is a geometric theory in the sense of D1.1.6 giving rise to an equivalent ($\mathfrak{BTop}/\mathbf{Set}$)-indexed category, and vice versa). We shall not give a formal proof of this equivalence; but in fact such a proof could be extracted from the list of examples of geometric theories which follows.

Examples 4.2.8 (*a*) The theory $\mathbb{M}$ of morphisms is a geometric theory, since it is a simple functional extension of the theory $\mathbb{O}_2$ of pairs of objects. More generally, for any finite category $\mathcal{D}$, the theory $\mathbb{D}$ of 4.2.4(*b*) may be obtained from the theory $\mathbb{O}_m$ (where m is the number of objects of $\mathcal{D}$) by making a finite number of simple functional extensions, followed by a finite number of simple geometric quotients to ensure that the commutativity of diagrams in $\mathcal{D}$ is respected (to force a diagram to commute, we force the equalizer of the two ways round it to be an isomorphism). So it is a geometric theory.

(*b*) More generally, for any finite family of finitary operation-symbols $\{\omega_1, \ldots, \omega_n\}$ (where each ω_i has a preassigned arity α_i which is a natural number), we may obtain the theory whose models are objects A equipped with morphisms $\omega_i: A^{\alpha_i} \to A$ as an iterated functional extension of $\mathbb{O}$, using 4.2.6(*b*). And we may then use iterated geometric quotients to impose equations between composites of these operations: that is, any finitely-presented (finitary) algebraic theory may be regarded as a geometric theory. (Examples of such theories include groups, rings, lattices, Heyting algebras (cf. A1.5.11), and so on.) And we may also consider many-sorted algebraic theories in this way, provided the number of sorts is finite.

(*c*) However, we may also consider geometric quotients corresponding to axioms which are not algebraic. For example, if our base topos $\mathcal{S}$ has a natural number object, then the theory of torsion groups (resp. torsion-free groups) is a geometric theory over $\mathcal{S}$, obtained from that of groups by adding the axiom

that $G_t \rightarrowtail G$ (resp. $e\colon 1 \to G_t$) is an isomorphism, where G_t is the geometric construct defined in 4.2.6(c).

(d) Let $\mathbb{T}$ be a theory over $\mathcal{S}$. By a *simple $\mathcal{S}$-specialization of* $\mathbb{T}$, we mean a theory $\mathbb{T}'$ such that, for some geometric construct F in $\mathbb{T}$ and some object I of $\mathcal{S}$, $\mathbb{T}'(\mathcal{E})$ is the subcategory of $\mathbb{T}(\mathcal{E})$ whose objects are those $\mathbb{T}$-models M for which $FM = p^*I$, and whose morphisms are those $\mathbb{T}$-model morphisms which induce the identity on p^*I. Thus, for example, the theory $\mathbb{O}_I$ of 4.2.4(d) is a simple $\mathcal{S}$-specialization of the theory $\mathbb{M}$ of morphisms. We claim that *if* $\mathbb{T}$ *is a geometric theory over* $\mathcal{S}$*, then so is any simple* $\mathcal{S}$*-specialization of* $\mathbb{T}$; for we may construct a theory equivalent to $\mathbb{T}'$ by first using a simple functional extension to adjoin a morphism of geometric constructs from F to the constant construct I, and then using a simple geometric quotient to make this morphism an isomorphism. Hence in particular $\mathbb{O}_I$ is a geometric theory for any object I of $\mathcal{S}$.

(e) The theory $\mathbb{S}\mathsf{ub}_I$ of 4.2.4(e) is a simple geometric quotient of $\mathbb{O}_I$, obtained by imposing on an $\mathbb{O}_I$-model $(A \to p^*I)$ the requirement that the diagonal map $A \rightarrowtail A \times_{p^*I} A$ should be an isomorphism. So it is a geometric theory.

(f) Generalizing (a), let $\mathbb{C}$ be an internal category in $\mathcal{S}$. Then the theory of diagrams of shape $\mathbb{C}$ is a geometric quotient of a simple functional extension of $\mathbb{O}_{C_0}$; so it is a geometric theory. (Alternatively, we could obtain this theory by first forming the theory whose models are pairs of internal categories with a discrete opfibration between them, and then 'specializing' the codomain of the opfibration to be the particular category $\mathbb{C}$.)

(g) Similarly, we may replace the finite presentability restrictions which appeared in (b) by '$\mathcal{S}$-presentability' using the notion of $\mathcal{S}$-specialization: that is, we may consider finitary algebraic theories whose sorts, generating operations and defining equations are all indexed by objects of the base topos $\mathcal{S}$, rather than by finite sets. (A good example, albeit a single-sorted one, is the theory of modules over an internal ring in $\mathcal{S}$.) And the same applies to the more complex theories considered in (c): forcing an 'I-indexed family' of simple geometric axioms to hold is the same as forcing a single morphism $F \to F'$ to be an isomorphism, where F and F' are 'I-indexed families of geometric constructs', that is, geometric constructs equipped with morphisms to the constant construct I.

(h) The theory of filtered categories is a geometric quotient of the theory of (arbitrary) categories, and hence a geometric theory: we simply have to add to the theory of categories three geometric axioms which say that the images of the three morphisms of 2.6.2(a) are isomorphisms. Combining this with (f) above, we deduce that the theory of $\mathbb{C}$-torsors is geometric, for any internal category $\mathbb{C}$ in $\mathcal{S}$.

(i) If $\mathcal{E}$ is a classifying topos for a geometric theory $\mathbb{T}$, then any sheaf subtopos $\mathbf{sh}_j(\mathcal{E})$ of $\mathcal{E}$ classifies a simple geometric quotient $\mathbb{T}_j$ of $\mathbb{T}$: for A4.3.11 tells us that a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ factors (uniquely) through $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$ iff f^* maps j-dense monomorphisms to isomorphisms, iff $f^*(d)$ is an isomorphism where $d\colon 1 \rightarrowtail J$ is the generic j-dense monomorphism. And, by the remarks

after 4.2.5, we may regard d as a morphism of geometric constructs in $\mathbb{T}$. Combining this observation with the last sentence of (h) above (and with Giraud's Theorem 3.3.4), we have shown that *for any bounded $\mathcal{S}$-topos $\mathcal{E}$, the theory $\mathfrak{BTop}/\mathcal{S}(\dot{-},\mathcal{E})$ classified by $\mathcal{E}$ is geometric.*

The main theorem of this section is that the result just stated has a converse, provided the base topos $\mathcal{S}$ has a natural number object.

Theorem 4.2.9 *Let $\mathcal{S}$ be a topos such that the classifying topos $\mathcal{S}[\mathbb{O}]$ exists in $\mathfrak{BTop}/\mathcal{S}$ (for example, any topos with a natural number object). Then, for any geometric theory $\mathbb{T}$ over $\mathcal{S}$, the classifying topos $\mathcal{S}[\mathbb{T}]$ exists.*

Proof Clearly, it suffices to prove (i) that the theory $\mathbb{O}_m$ of m-tuples of objects has a classifying topos, (ii) that if $\mathbb{T}$ has a classifying topos then so does any simple functional extension of $\mathbb{T}$, and (iii) that if $\mathbb{T}$ has a classifying topos then so does any simple geometric quotient of $\mathbb{T}$. But each of these is easy, given the results of the previous section. For (i), we take the m-fold product of copies of the object classifier $\mathcal{S}[\mathbb{O}]$ in $\mathfrak{BTop}/\mathcal{S}$. For (ii), suppose we wish to adjoin a morphism $F \to F'$, where F and F' are geometric constructs in $\mathbb{T}$. We may regard F and F' as geometric morphisms

$$\mathcal{S}[\mathbb{T}] \underset{f'}{\overset{f}{\rightrightarrows}} \mathcal{S}[\mathbb{O}]$$

in $\mathfrak{BTop}/\mathcal{S}$; and it is then clear that the inserter of this parallel pair of morphisms, as defined in 1.1.4(d), is a classifying topos for the extended theory $\mathbb{T}'$. Similarly for (iii), if $\mathbb{T}'$ is a simple geometric quotient of $\mathbb{T}$ then we may construct its classifying topos as the inverter (1.1.4(e)) of a 2-cell between two morphisms $\mathcal{S}[\mathbb{T}] \rightrightarrows \mathcal{S}[\mathbb{O}]$. □

We have seen that the only step in the construction of classifying toposes which requires a natural number object in the base topos is the initial one of constructing an object classifier; the rest of the construction is insensitive to the particular properties of the base. It is therefore of interest to ask whether an object classifier might exist over some topos without a natural number object. However, the answer is no: the following argument is due to A. Blass [130].

Lemma 4.2.10 *Let $\mathbb{T}$ be a finitely-presented algebraic theory, and suppose there exists a classifying topos $p: \mathcal{S}[\mathbb{T}] \to \mathcal{S}$ for $\mathbb{T}$ over $\mathcal{S}$. Then $\mathbb{T}(\mathcal{S})$ has an initial object.*

Proof Consider $p_*(G)$, where G is the generic $\mathbb{T}$-model in $\mathcal{S}[\mathbb{T}]$. Since p_* is cartesian, $p_*(G)$ has a $\mathbb{T}$-model structure in $\mathcal{S}$. But given any $\mathbb{T}$-model A in $\mathcal{S}$, there exists a geometric morphism $f: \mathcal{S} \to \mathcal{S}[\mathbb{T}]$ over $\mathcal{S}$ such that $f^*(G) \cong A$,

and hence we have a composite morphism

$$p_*(G) \xrightarrow{\sim} f^*p^*p_*(G) \xrightarrow{f^*(\epsilon_G)} f^*(G) \xrightarrow{\sim} A$$

in $\mathbb{T}(\mathcal{S})$, where ϵ is the counit of $(p^* \dashv p_*)$.

Thus we have shown that $p_*(G)$ is a weakly initial object in $\mathbb{T}(\mathcal{S})$, i.e. that there exists at least one morphism from it to any other $\mathbb{T}$-model in $\mathcal{S}$. In fact, as will be clear from the construction of classifying toposes given in Section D3.1, $p_*(G)$ is actually initial in $\mathbb{T}(\mathcal{S})$; but it does not seem possible to prove this directly. However, it is clear that for any $\mathbb{T}$-model A in $\mathcal{S}$ we may form the subobject $P_{\mathbb{T}}A$ of PA which is 'the object of sub-$\mathbb{T}$-models of A', by intersecting a finite number of equalizers, and that this subobject is closed under internal intersections in PA; hence we may obtain the unique smallest subobject $A' \rightarrowtail A$ which is a sub-$\mathbb{T}$-model of A. If we apply this to $p_*(G)$, the sub-$\mathbb{T}$-model A' we obtain is still weakly initial in $\mathbb{T}(\mathcal{S})$; but, for any two $\mathbb{T}$-model homomorphisms $A' \rightrightarrows B$, their equalizer is a sub-$\mathbb{T}$-model of A', and so must be the whole of A'. Hence A' is actually an initial object. □

Theorem 4.2.11 *For a topos $\mathcal{S}$, the following are equivalent:*

(i) *$\mathcal{S}$ has a natural number object.*

(ii) *$\mathfrak{BTop}/\mathcal{S}$ has an object classifier.*

(iii) *$\mathfrak{BTop}/\mathcal{S}$ has classifiers for all geometric theories over $\mathcal{S}$.*

(iv) *$\mathfrak{BTop}/\mathcal{S}$ has a classifier for the algebraic theory freely generated by one unary and one nullary operation.*

Proof (i) $\Rightarrow$ (ii) was proved in 3.2.9, and (ii) $\Rightarrow$ (iii) in 4.2.9. (iii) $\Rightarrow$ (iv) follows from 4.2.8(b), and (iv) $\Rightarrow$ (i) follows from 4.2.10, since a natural number object is by definition an initial model of the theory described in (iv) (cf. A2.5.1). □

Nevertheless, over a topos $\mathcal{S}$ without a natural number object we may construct classifying toposes for all *propositional* geometric theories. The formal definition of a propositional geometric theory will be found in D1.1.7(m); for present purposes, it may be taken to be theory built up by an inductive process like that of 4.2.7, but starting from the theory of a finite family of subobjects of (given) objects of the base topos rather than from the theory of m-tuples of arbitrary objects.

Proposition 4.2.12 *For any topos $\mathcal{S}$, every propositional geometric theory $\mathbb{T}$ over $\mathcal{S}$ has a classifying topos $\mathcal{S}[\mathbb{T}]$, which is localic over $\mathcal{S}$.*

Proof The outline of the argument is similar to that of 4.2.9, but we take 3.2.10 rather than 3.2.9 as our starting-point. That is, if our initial theory $\mathbb{T}_0$ is the theory of m-tuples $(A_1 \rightarrowtail I_1, A_2 \rightarrowtail I_2, \ldots, A_m \rightarrowtail I_m)$ (where $I_1, I_2, \ldots, I_m$ are given objects of $\mathcal{S}$), we take $\mathcal{S}[\mathbb{T}_0]$ to be the product in $\mathfrak{BTop}/\mathcal{S}$ of the

diagram toposes $[\mathbb{K}(I_j), \mathcal{S}]$ $(1 \leq j \leq m)$. For the inductive steps, we cannot argue from the existence of finite weighted limits in $\mathfrak{BTop}/\mathcal{S}$ as we did in the proof of 4.2.9, since we no longer have the object classifier available; but we can argue directly as follows. If $\mathbb{T}'$ is a simple geometric quotient of $\mathbb{T}$, corresponding to the requirement that $f: F \to F'$ be invertible, then we can take $\mathcal{S}[\mathbb{T}']$ to be $\mathbf{sh}_j(\mathcal{S}[\mathbb{T}])$, where j is the smallest local operator whose associated sheaf functor maps $f_G: FG \to F'G$ (where G is the generic $\mathbb{T}$-model) to an isomorphism (cf. A4.5.14(c)). For simple functional extensions, we note first that the diagram category $[\mathbb{K}(FG), \mathcal{S}[\mathbb{T}]]$ classifies $\mathbb{T}$-models M equipped with a distinguished subobject of FM; so by first adjoining a generic subobject of $F \times F'$ in this way, and then using a simple geometric quotient to force it to be the graph of a morphism $F \to F'$, we may obtain a classifying topos for a simple functional extension of $\mathbb{T}$. Finally, we observe that the two constructions we have used (taking sheaves for a local operator, and taking diagrams on an internal poset) both give rise to localic geometric morphisms, by A4.6.2(a) and (d); since a composite of localic morphisms is localic by A4.6.2(e), it follows that any classifying topos constructed in this way will be localic over $\mathcal{S}$. □

Once again, Proposition 4.2.12 has a converse: for any localic $\mathcal{S}$-topos $\mathcal{E}$, the theory $\mathbb{T}$ defined by $\mathbb{T}(\mathcal{F}) = \mathfrak{BTop}/\mathcal{S}\,(\mathcal{F}, \mathcal{E})$ is (equivalent to) a propositional geometric theory over $\mathcal{S}$. The proof of this follows from the representation of localic toposes in 3.3.5, together with the observation in 3.2.4(d) that if $\mathbb{P}$ is a preorder then the theory of $\mathbb{P}$-torsors is equivalent to the theory of ideals of $\mathbb{P}$ (which is easily seen to be a propositional geometric theory). Cf. also D1.2.15(m).

Suggestions for further reading: Blass [130], Tierney [1169].

B4.3 Some exponentiable toposes

The 2-category $\mathfrak{BTop}/\mathcal{S}$ is not cartesian closed, but it contains an important class of exponentiable objects, i.e. $\mathcal{S}$-toposes $p: \mathcal{E} \to \mathcal{S}$ such that the 2-functor $(-) \times_{\mathcal{S}} \mathcal{E}: \mathfrak{BTop}/\mathcal{S} \to \mathfrak{BTop}/\mathcal{S}$ has a right 2-adjoint. The precise determination of this class belongs properly to Part C of this book, since it relies on a topologically-inspired notion of 'local compactness' for $\mathcal{S}$-toposes; and we shall describe it in Section C4.4. In the present section, we shall see what can be said about exponentials in $\mathfrak{BTop}/\mathcal{S}$ by the general 2-categorical methods of this chapter. We shall make heavy use of the theory of classifying toposes developed in the previous section, and in particular of the object classifier $\mathcal{S}[\mathbb{O}]$; therefore we shall assume throughout this section that our base topos $\mathcal{S}$ has a natural number object.

We say that a particular exponential $\mathcal{F}^{\mathcal{E}}$ exists if the functor

$$\mathfrak{BTop}/\mathcal{S}\,((-) \times_{\mathcal{S}} \mathcal{E}, \mathcal{F}): \mathfrak{BTop}/\mathcal{S}^{\mathrm{op}} \longrightarrow \mathfrak{CAT}$$

is representable (as usual, we mean this in the 'up-to-equivalence' sense).

Theorem 4.3.1 *Let $\mathcal{S}$ be a topos with a natural number object. Then a bounded $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ is exponentiable iff the particular exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$ exists.*

Proof We rely on the fact, established in 4.2.8(i), that every bounded $\mathcal{S}$-topos occurs as the classifying topos for some geometric theory: what we shall actually prove is that if $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$ exists then so does $\mathcal{S}[\mathbb{T}]^{\mathcal{E}}$ for any geometric theory $\mathbb{T}$ over $\mathcal{S}$. Now in the proof of 4.2.9 we saw how all the toposes $\mathcal{S}[\mathbb{T}]$ may be built up by finite weighted limits from $\mathcal{S}[\mathbb{O}]$; and since $(-)^{\mathcal{E}}$, to the extent that it exists, is a right adjoint, it should preserve these weighted limits. Thus, for example, to construct the classifying topos for a simple functional extension $\mathbb{T}'$ of $\mathbb{T}$, we formed the inserter of a diagram of the form

$$\mathcal{S}[\mathbb{T}] \underset{f'}{\overset{f}{\rightrightarrows}} \mathcal{S}[\mathbb{O}];$$

so in order to construct $\mathcal{S}[\mathbb{T}']^{\mathcal{E}}$ from $\mathcal{S}[\mathbb{T}]^{\mathcal{E}}$, we simply form the inserter of the diagram

$$\mathcal{S}[\mathbb{T}]^{\mathcal{E}} \underset{f'^{\mathcal{E}}}{\overset{f^{\mathcal{E}}}{\rightrightarrows}} \mathcal{S}[\mathbb{O}]^{\mathcal{E}}$$

and verify that it has the desired universal property. (Here $f^{\mathcal{E}}$ of course corresponds to the composite

$$\mathcal{S}[\mathbb{T}]^{\mathcal{E}} \times_{\mathcal{S}} \mathcal{E} \xrightarrow{\text{ev}} \mathcal{S}[\mathbb{T}] \xrightarrow{f} \mathcal{S}[\mathbb{O}],$$

where ev corresponds to the identity morphism on $\mathcal{S}[\mathbb{T}]^{\mathcal{E}}$.) □

Another way of stating 4.3.1 involves a notion of 'duality' for geometric theories. If $\mathbb{T}$ is a geometric theory over $\mathcal{S}$, then for any bounded $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ we may also regard it (by restriction) as a geometric theory over $\mathcal{E}$, and a purely formal argument shows that there is a pullback square

$$\begin{tikzcd} \mathcal{E}[\mathbb{T}] \arrow[r] \arrow[d] & \mathcal{S}[\mathbb{T}] \arrow[d] \\ \mathcal{E} \arrow[r, "p"] & \mathcal{S} \end{tikzcd}$$

in $\mathfrak{BTop}$. Thus we may regard the assignment $\mathcal{E} \mapsto \mathcal{E}[\mathbb{T}]$ as itself defining a theory $\mathbb{T}^*$ over $\mathcal{S}$ in the sense of 4.2.1 (the effect of $\mathbb{T}^*$ on a geometric morphism f being given by the inverse image of the pullback of f along $\mathcal{S}[\mathbb{T}] \to \mathcal{S}$, and its effect on geometric transformations being obtained by the method of 3.2.14). We say

that the theory $\mathbb{T}$ is *dualizable* if $\mathbb{T}^*$ is itself a geometric theory (equivalently, has a classifying topos). (The word 'dual' should not be taken too seriously in this context: the relation between $\mathbb{T}$ and $\mathbb{T}^*$ is far from symmetric. The paradigm one should have in mind is that of the dual of an infinite-dimensional vector space. However, we note that the construction is indeed contravariant: given a morphism of theories $\mathbb{S} \to \mathbb{T}$, that is a family of functors $\mathbb{S}(\mathcal{E}) \to \mathbb{T}(\mathcal{E})$ which are natural in $\mathcal{E}$, we obtain a geometric morphism $\mathcal{E}[\mathbb{S}] \to \mathcal{E}[\mathbb{T}]$ for each $\mathcal{E}$, classifying the $\mathbb{T}$-model obtained by applying the given functor to the generic $\mathbb{S}$-model, and the inverse images of these morphisms define a morphism of theories $\mathbb{T}^* \to \mathbb{S}^*$.)

Corollary 4.3.2 *A bounded $\mathcal{S}$-topos is exponentiable iff it is the classifying topos over $\mathcal{S}$ of a dualizable geometric theory.*

Proof Suppose $\mathcal{E} \simeq \mathcal{S}[\mathbb{T}]$ and $\mathbb{T}^*$ is geometric. Then $\mathcal{S}[\mathbb{T}^*]$ has the universal property of an exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$; for we have

$$\begin{aligned}\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{S}[\mathbb{T}^*]) &\simeq \mathbb{T}^*(\mathcal{F})\\ &= \mathcal{F}[\mathbb{T}]\\ &\simeq \mathcal{F}\times_{\mathcal{S}}\mathcal{S}[\mathbb{T}]\\ &\simeq \mathfrak{BTop}/\mathcal{S}\,(\mathcal{F}\times_{\mathcal{S}}\mathcal{E},\mathcal{S}[\mathbb{O}])\end{aligned}$$

for any bounded $\mathcal{S}$-topos $\mathcal{F}$. Hence by 4.3.1 $\mathcal{E}$ is exponentiable. Conversely, if $\mathcal{E}$ is exponentiable and classifies a geometric theory $\mathbb{T}$, then the reverse of the above argument shows that the exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$ is a classifying topos for $\mathbb{T}^*$. □

We already know one class of examples of dualizable geometric theories: namely, the theories of $\mathbb{C}$-torsors, for arbitrary internal categories $\mathbb{C}$ in $\mathcal{S}$.

Lemma 4.3.3 *For any internal category $\mathbb{C}$ in $\mathcal{S}$, the diagram category $[\mathbb{C},\mathcal{S}]$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$.*

Proof By 3.2.14 we have $\mathcal{E}\times_{\mathcal{S}}[\mathbb{C},\mathcal{S}] \simeq [p^*\mathbb{C},\mathcal{E}]$ for any $p\colon \mathcal{E}\to\mathcal{S}$, so if $\mathbb{T}$ is the theory of $\mathbb{C}$-torsors then we can take $\mathbb{T}^*$ to be the theory of diagrams of shape $\mathbb{C}$ in $\mathcal{S}$-toposes – recall that we saw in 4.2.8(f) that this is a geometric theory. □

Remark 4.3.4 As a special case of 4.3.3, we note that the slice category $\mathcal{S}/I$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, for any object I of $\mathcal{S}$. It follows that $\mathfrak{BTop}/\mathcal{S}$ has '$\mathcal{S}$-indexed products', as we claimed in Section B4.1: that is, if we are given an I-indexed family of bounded $\mathcal{S}$-toposes (that is, a bounded $\mathcal{S}/I$-topos $p\colon \mathcal{E}\to\mathcal{S}/I$), we may form a bounded $\mathcal{S}$-topos $\Pi_I(\mathcal{E})$ such that, for any bounded $q\colon \mathcal{F}\to\mathcal{S}$, we have

$$\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\Pi_I(\mathcal{E})) \simeq \mathfrak{BTop}/(\mathcal{S}/I)\,(\mathcal{F}/q^*I,\mathcal{E})\,.$$

To do this, we simply form the pullback

$$\begin{tikzcd} \Pi_I(\mathcal{E}) \arrow[r] \arrow[d] & \mathcal{E}^{\mathcal{S}/I} \arrow[d] \\ \mathcal{S} \arrow[r] & \mathcal{S}/I^{\mathcal{S}/I} \end{tikzcd}$$

as in A1.5.2(i).

Next, we note three basic 'stability properties' of exponentiable toposes.

Lemma 4.3.5

(i) *Exponentiability is stable under composition; that is, if $p: \mathcal{E} \to \mathcal{S}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, then so is $pq: \mathcal{F} \to \mathcal{S}$ for any exponentiable object $q: \mathcal{F} \to \mathcal{E}$ of $\mathfrak{BTop}/\mathcal{E}$.*

(ii) *Exponentiability is stable under pullback; that is, if $p: \mathcal{E} \to \mathcal{S}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$ and we have a pullback*

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[d, "q"] & \mathcal{E} \arrow[d, "p"] \\ \mathcal{T} \arrow[r, "f"] & \mathcal{S} \end{tikzcd}$$

then $q: \mathcal{F} \to \mathcal{T}$ is exponentiable in $\mathfrak{BTop}/\mathcal{T}$.

(iii) *Exponentiability is stable under retracts; that is, if $q: \mathcal{F} \to \mathcal{S}$ is a retract in $\mathfrak{BTop}/\mathcal{S}$ of an exponentiable topos $p: \mathcal{E} \to \mathcal{S}$, then q is exponentiable.*

Proof (i) Generalizing the observation in 4.3.4, an object $p: \mathcal{E} \to \mathcal{S}$ is exponentiable iff the pullback functor $p^*: \mathfrak{BTop}/\mathcal{S} \to \mathfrak{BTop}/\mathcal{E}$ has a right adjoint. The latter condition is clearly stable under composition; so the result follows immediately.

(ii) Suppose $\mathcal{E}$ classifies a geometric theory $\mathbb{T}$ over $\mathcal{S}$. Then it is clear from the definition of the dual theory $\mathbb{T}^*$ that its restriction to a theory over $\mathcal{T}$ is the dual of the corresponding restriction of $\mathbb{T}$; so it is geometric if $\mathbb{T}^*$ is.

(iii) is immediate from the fact that $\mathcal{S}[\mathbb{O}]^{(-)}$, to the extent that it is defined, is contravariantly functorial; if e is the idempotent endomorphism of $\mathcal{E}$ corresponding to the retract $\mathcal{F}$, then $\mathcal{S}[\mathbb{O}]^e$ is an idempotent endomorphism of $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$, and if we split it (cf. 1.1.9(*b*)) we obtain a topos having the universal property of $\mathcal{S}[\mathbb{O}]^{\mathcal{F}}$. □

Combining 4.3.3 and 4.3.5(i), we have

Corollary 4.3.6 *If $\mathcal{E}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, then so is $[\mathbb{C}, \mathcal{E}]$ for any internal category $\mathbb{C}$ in $\mathcal{E}$.* □

In particular, any slice of an exponentiable topos is exponentiable. In the converse direction, we have

Lemma 4.3.7 *If $\mathcal{E}/A$ is exponentiable for some well-supported object A of $\mathcal{E}$, then $\mathcal{E}$ is exponentiable.*

Proof If $\mathcal{E}/A$ is exponentiable, then so is $\mathcal{E}/A^n$ for any natural number n, by the remark immediately before the statement of the lemma. But by 1.5.6(iii) we may identify $\mathcal{E}$ with the category of objects of $\mathcal{E}/A$ equipped with descent data relative to the morphism $A \twoheadrightarrow 1$; equivalently, by 3.4.12, with the (pseudo-) colimit of the simplicial topos $\mathcal{E}/A^\bullet$, or of its 2-truncation. Hence we may define $\mathcal{F}^\mathcal{E}$, for any bounded $\mathcal{S}$-topos $\mathcal{F}$, to be the pseudo-limit of the 2-truncation of the cosimplicial topos $\mathcal{F}^{\mathcal{E}/A^\bullet}$. □

In contrast to 3.4.7, the property of exponentiability does not 'descend' along more general descent morphisms in $\mathfrak{BTop}/\mathcal{S}$ (see C4.4.9).

In addition to diagram categories $[\mathbb{C}, \mathcal{S}]$, the other class of bounded $\mathcal{S}$-toposes whose exponentiability we shall need to invoke in the next section is the class of coherent $\mathcal{S}$-toposes. Unfortunately, it seems hard to give a precise definition of this class, let alone to prove that its members are exponentiable, without trespassing on material which properly belongs in Parts C and D of this book; so the reader who has not yet mastered those parts will have to be content for the moment with a rather informal account. (Such a reader can rest assured that we shall give rigorous definitions of the notions of coherent theory and of coherent topos in Sections D1.1 and D3.3 respectively, and that a rigorous proof of the exponentiability of coherent toposes will follow easily from the results of those sections and of Section C4.4 – see in particular C4.4.6(a).)

Informally, then, we say that a geometric construct F in a theory $\mathbb{T}$, in the sense of 4.2.5, is *coherent* if $F(M)$ can be built up from the objects defining the $\mathbb{T}$-model M using only those categorical constructions (finite limits, finite coproducts and coequalizers of equivalence relations) which are preserved by arbitrary coherent functors (and not just inverse image functors). And we say that $\mathbb{T}$ is a *coherent theory* if it can be obtained from the theory of diagrams of shape $\mathbb{C}$, for some internal category $\mathbb{C}$ in $\mathcal{S}$, by adjoining '$\mathcal{S}$-indexed families of coherent axioms', that is families of axioms which assert that morphisms between two coherent constructs are isomorphic (cf. 4.2.7).

For example, any finitely-presented algebraic theory, as considered in 4.2.8(b), is coherent. On the other hand, the theory of torsion groups, as defined in 4.2.8(c), is not coherent, because the torsion subgroup G_t is not a coherent construct (being defined as an infinite union). But the theory of torsion-free groups *is* coherent, because instead of imposing the single geometric axiom which says that $1 \to G_t$ is an isomorphism, we can impose the N-indexed family of axioms which say that $1 \to G_n$ is an isomorphism (using the notation of 4.2.6(c)), and each G_n is a coherent construct. Rather more subtly, the assertion that a given diagram of shape $\mathbb{C}^{\mathrm{op}}$ is a $\mathbb{C}$-torsor cannot, in general, be expressed by coherent axioms – even though the theory of filtered categories is coherent.

This is because, for example, the first condition of 2.6.2(a) for filteredness of the total category of the discrete fibration $\mathbb{D} \to \mathbb{C}$ corresponding to a diagram D involves consideration of the geometric construct D_0, which is not coherent (unless C_0 is finite) because it is the C_0-indexed coproduct of the objects in the diagram. However, if $\mathbb{C}$ is finite then the notion of $\mathbb{C}$-torsor is coherent; also, if $\mathbb{C}$ is cocartesian, then we know that $\mathbb{C}$-torsors are the same thing as cartesian functors defined on $\mathbb{C}^{\mathrm{op}}$, by 3.2.5, and these can clearly be defined by coherent axioms. (For further examples of what can and cannot be expressed by coherent axioms, see D1.1.7 and the remarks at the end of Section D2.4.)

Of course, we say a bounded $\mathcal{S}$-topos is coherent if it is (equivalent to) the classifying topos of a coherent theory over $\mathcal{S}$. The key result which we require about coherent toposes is

Theorem 4.3.8 *A bounded $\mathcal{S}$-topos is coherent iff it can be presented as the topos $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ of $\mathcal{S}$-valued sheaves on an internal site $(\mathbb{C}, T)$ in $\mathcal{S}$ such that the category $\mathbb{C}$ is cartesian and the coverage T consists entirely of finite families (that is, families indexed by finite cardinals in $\mathcal{S}$).*

Proof See D3.3.1 for a detailed proof. Intuitively, the reason why the result is true is that the classifying topos for the theory of diagrams of shape $\mathbb{D}$ may be identified with a topos of the form $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ where $\mathbb{C}$ is cartesian (specifically, $\mathbb{C}^{\mathrm{op}}$ is the category of all diagrams of shape $\mathbb{D}$ which are finite colimits of representables – we may construct this as an internal full subcategory of $[\mathbb{D}, \mathcal{S}]$), and imposing a coherent axiom on a diagram F amounts to saying that the functor $[\mathbb{D}, \mathcal{S}]\,(-, F) \colon \mathbb{C}^{\mathrm{op}} \to \mathcal{S}$ carries certain finite families in $\mathbb{C}^{\mathrm{op}}$ to epimorphic families. Thus if we impose the coverage on $\mathbb{C}$ in which all such families cover, we obtain a classifying topos for diagrams satisfying the given coherent axioms. $\square$

We shall also require the following result from Part C:

Theorem 4.3.9 *Suppose given an internal site $(\mathbb{C}, T)$ in a topos $\mathcal{S}$, and a geometric morphism $p \colon \mathcal{E} \to \mathcal{S}$. Then there exists an internal coverage $p^{\#}T$ on $p^*\mathbb{C}$ in $\mathcal{E}$, such that the square*

$$\begin{tikzcd}
\mathbf{Sh}_{\mathcal{E}}(p^*\mathbb{C}, p^{\#}T) \arrow[r] \arrow[d] & \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T) \arrow[d] \\
\mathcal{E} \arrow[r, "p"] & \mathcal{S}
\end{tikzcd}$$

is a pullback. Moreover, if $\mathbb{C}$ is cartesian and all covers in T are finite, then all covers in $p^{\#}T$ are finite; and this coverage is such that, under the correspondence of 2.3.14 between diagrams of shape $p^*\mathbb{C}^{\mathrm{op}}$ in $\mathcal{E}$ and $\mathcal{S}$-indexed functors $[\mathbb{C}^{\mathrm{op}}] \to \mathbb{E}$, the $p^{\#}T$-sheaves in $\mathcal{E}$ correspond to the functors $[\mathbb{C}^{\mathrm{op}}] \to \mathbb{E}$ which satisfy the sheaf axiom for covers in T.

Proof The first assertion of the theorem is more or less immediate from the proof of 3.3.6, given that every subtopos of $[p^*\mathbb{C}^{\mathrm{op}}, \mathcal{E}]$ can be written in the form $\mathbf{Sh}_{\mathcal{E}}(p^*\mathbb{C}, R)$ for some coverage R. However, in order to establish the second assertion, we need an explicit description of the 'pullback coverage' $p^{\#}T$. This is provided in Section C2.4, and the result claimed here follows directly from it (see C2.4.7). □

Given these two results, it becomes a triviality to deduce

Corollary 4.3.10 *Any coherent theory $\mathbb{T}$ over $\mathcal{S}$ is dualizable; equivalently, any coherent $\mathcal{S}$-topos is exponentiable. Moreover, the dual theory $\mathbb{T}^*$ is also coherent.*

Proof If $\mathbb{T}$ is classified by a topos of the form $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ as in 4.3.8, then the dual theory $\mathbb{T}^*$ is simply the theory of T-sheaves on $\mathbb{C}$, by 4.3.9. And this is clearly geometric, and indeed coherent, because the assertion that a functor F defined on $\mathbb{C}^{\mathrm{op}}$ satisfies the sheaf axiom for a finite covering family $(U_i \to U \mid 1 \leq i \leq n)$ is just the assertion that $F(U)$ maps isomorphically to the limit of a certain finite diagram constructed from F (cf. the remarks after A2.1.9). So, in order to say that F is a T-sheaf, we have simply to impose a T-indexed family of such coherent axioms. □

As a particular case of 4.3.10, we note that if $\mathbb{C}$ is a coherent internal category in $\mathcal{S}$ and T is its coherent coverage (defined as in A2.1.11(*b*)), then the topos $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ is exponentiable. (In fact this 'particular case' includes all coherent $\mathcal{S}$-toposes; cf. D3.3.1.)

Suggestion for further reading: Johnstone & Joyal [542].

B4.4 Fibrations and partial products

One problem with slice categories of a 2-category, which we have already encountered in 3.2.14 above, is that in passing from a 2-category $\mathfrak{K}$ to a slice $\mathfrak{K}/A$ we 'lose some of the 2-cells'; that is, if $f, g\colon (p\colon B \to A) \rightrightarrows (q\colon C \to A)$ are two morphisms of $\mathfrak{K}/A$, not every 2-cell $\alpha\colon f \to g$ in $\mathfrak{K}$ is a 2-cell of $\mathfrak{K}/A$. In particular, if $\mathfrak{K}$ has (pseudo-)pullbacks, the operation of pulling back along a fixed morphism $p\colon B \to A$ on $\mathfrak{K}$ is defined on 1-cells $q\colon C \to A$, but not on arbitrary 2-cells with codomain A.

As we saw in 3.2.14, the presence of tensors with $\mathbf{2}$ in $\mathfrak{K}$ can be used to circumvent this problem in some cases. However, there is another possible way round, using an intrinsic 2-categorical notion of fibration; we devote the first part of this section to setting it up, and to investigating the fibrations and opfibrations in the particular 2-category $\mathfrak{BTop}/\mathcal{S}$.

To motivate the definition, let us consider what happens when we pull back arbitrary functors $F\colon \mathcal{D} \to \mathcal{C}$ along a fibration $\Pi\colon \mathcal{F} \to \mathcal{C}$ in $\mathfrak{CAT}$. We recall from 1.3.9 that in this case the pseudo-pullback $\mathcal{D} \times_{\mathcal{C}} \mathcal{F}$ (whose objects are triples

(B, E, f) where $B \in \text{ob}\,\mathcal{D}$, $E \in \text{ob}\,\mathcal{F}$ and f is an isomorphism $F(B) \to \Pi(E)$ in $\mathcal{C}$) is equivalent to the strict pullback (the full subcategory of objects for which f is an identity morphism). We shall exploit this by working in what follows with strict pullbacks rather than pseudo-pullbacks.

Now suppose we are given a natural transformation $\alpha\colon F \to G$ between functors $\mathcal{D} \rightrightarrows \mathcal{C}$. The pullbacks $\Pi^*(F)$ and $\Pi^*(G)$ are not in general equivalent in $\mathfrak{Cat}/\mathcal{F}$; but we have a canonical functor from the second to the first, which sends an object (B, E) with $G(B) = \Pi(E)$ to $(B, \text{dom}\,\widetilde{\alpha_B})$ where $\widetilde{\alpha_B}$ is a prone lifting of $\alpha_B\colon F(B) \to G(B)$ with codomain E. Moreover, the prone lifting $\widetilde{\alpha_B}$ itself defines a natural transformation fitting into the right-hand cell of the diagram

$$\begin{tikzcd}
 & F^*\mathcal{F} \arrow[r] \arrow[dl] & \mathcal{F} \arrow[dl, phantom, "\downarrow"] \\
\mathcal{D} & G^*\mathcal{F} \arrow[l] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

of which the left-hand cell commutes (up to isomorphism).

Abstracting from this example, we introduce the following definition.

Definition 4.4.1 Let $\mathfrak{K}$ be a 2-category with (pseudo-)pullbacks. We say a morphism $p\colon E \to B$ of $\mathfrak{K}$ is a *fibration* if we are given operations assigning to any 2-cell $\alpha\colon f \to g$ between 1-cells $f, g\colon A \rightrightarrows B$ a 1-cell $l(\alpha)\colon g^*(E) \to f^*(E)$ together with a 2-isomorphism $f^*(p) \circ l(\alpha) \cong g^*(p)$ and a (not necessarily invertible) 2-cell $\tilde{\alpha}\colon p^*(f) \circ l(\alpha) \to p^*(g)$, which 'lies over α' in the sense that the 2-cells obtained by pasting the diagrams

$$\begin{tikzcd}
g^*(E) \arrow[r, "l(\alpha)"] \arrow[dr] \arrow[dd] \arrow[r, phantom, shift right=6, "\downarrow\tilde{\alpha}"] & f^*(E) \arrow[d] \\
 & E \arrow[d] \\
A \arrow[r, "g"] & B
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
g^*(E) \arrow[r, "l(\alpha)"] \arrow[d] & f^*(E) \arrow[dl] \arrow[d] \\
A \arrow[d, equal] \arrow[dr, "f"] \arrow[d, phantom, shift left=8, "\alpha\downarrow"] & E \arrow[d] \\
A \arrow[r, "g"'] & B
\end{tikzcd}$$

are equal. These data are required to satisfy the following conditions:

(i) They are compatible with vertical composition of 2-cells, in the sense that $l(1_f) = 1_{f^*E}$ and $\widetilde{1_f}$ is the identity 2-cell, and if $\beta\colon g \to h$ is another 2-cell then we have an isomorphism $l(\alpha)\circ l(\beta) \cong l(\beta\alpha)$, compatible in the obvious sense with the isomorphisms over A and the 2-cells with codomain E.

(ii) They are also compatible with horizontal composition of a 2-cell with a 1-cell, in the sense that if we are given $k: C \to A$ then $l(\alpha \circ k)$ is the pullback of $l(\alpha)$ along k, and $\widetilde{\alpha \circ k}$ is obtained by pasting $\tilde{\alpha}$ with the canonical isomorphism in the pullback square

$$\begin{tikzcd} k^*g^*(E) \arrow[r] \arrow[d, "l(\alpha \circ k)"] & g^*(E) \arrow[d, "l(\alpha)"] \\ k^*f^*(E) \arrow[r] & f^*(E) \end{tikzcd} .$$

(iii) Given any 1-cell $m: D \to E$ together with 1-cells $x: D \to f^*E$, $y: D \to g^*E$ such that $(p^*g)y \cong m$, $(f^*p)x \cong (g^*p)y$ and we are given a 2-cell $\beta: (p^*f)x \to m$, there exists a unique 2-cell $\gamma: x \to l(\alpha)y$ such that the composite

$$(f^*p)x \xrightarrow{(f^*p)\circ\gamma} (f^*p)l(\alpha)y \xrightarrow{\tilde{\alpha}\circ y} (g^*p)y$$

is the given isomorphism and $(p^*f) \circ \gamma$ is similarly isomorphic to β.

Dually, we say p is an *opfibration* if it satisfies the same properties as above, but with the direction of the 1-cell $l(\alpha)$ reversed.

The definitions may alternatively be phrased in terms of the action of p on suitable 'lax slice 2-categories' of $\mathfrak{K}$; this is done in [536], where it is shown that the structure making a morphism into a fibration is unique up to canonical isomorphism if it exists. However, for our present purposes it seems more helpful to spell out the meaning in elementary terms, as above.

Example 4.4.2 Suppose $\mathfrak{K}$ has cotensors with $\mathbf{2}$ as well as pullbacks, and let

$$\begin{tikzcd} (f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\to"] & C \arrow[d, "g"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

be a comma square in $\mathfrak{K}$. We claim that p *has a fibration structure* (and, dually, q has an opfibration structure). To see this, suppose given a 2-cell $\alpha: h \to k$ between 1-cells $D \rightrightarrows A$. By the construction of 1.1.4(c), the pullbacks $h^*(f \downarrow g)$ and $k^*(f \downarrow g)$ may be identified with the comma objects $(fh \downarrow g)$ and $(fk \downarrow g)$ respectively. Let p_h, q_h and p_k, q_k respectively denote the projections from the two pullbacks to D and to $(f \downarrow g)$; then we may obtain a further

morphism $m: k^*(f \downarrow g) \to (f \downarrow g)$ from the morphisms hp_k and qq_k, together with the composite 2-cell

$$fhp_k \xrightarrow{f \circ \alpha \circ p_k} fkp_k \cong fpq_k \longrightarrow gqq_k \,.$$

Then the pair (p_k, m) induces a morphism $l(\alpha): k^*(f \downarrow g) \to h^*(f \downarrow g)$; and we also have a 2-cell $\tilde{\alpha}$ from $q_h l(\alpha) \cong m$ to q_k, induced in the obvious way by α. It is straightforward to verify that these constructions of $l(\alpha)$ and $\tilde{\alpha}$ satisfy the conditions of 4.4.1

Lemma 4.4.3

(i) *A composite of fibrations is a fibration.*

(ii) *A pullback of a fibration is a fibration.*

Proof (i) Suppose given fibrations $p: E \to B$ and $q: F \to E$, and a 2-cell $\alpha: f \to g$ between morphisms $f, g: A \rightrightarrows B$. We may first construct $l(\alpha): g^*E \to f^*E$ and the 2-cell $\tilde{\alpha}: p^*(f) \circ l(\alpha) \to p^*(g)$; then the latter induces a 1-cell $l'(\tilde{\alpha})$ and a 2-cell $\tilde{\tilde{\alpha}}$ from the composite $q^*p^*(f) \circ q^*(l(\alpha)) \circ l'(\tilde{\alpha})$ to $q^*p^*(g)$. It is then straightforward to verify that the composite $q^*(l(\alpha)) \circ l'(\tilde{\alpha})$ and the 2-cell $\tilde{\tilde{\alpha}}$ provide the structure needed to make the composite pq into a fibration.

(ii) Suppose given a pullback square

$$\begin{tikzcd}
F \arrow[r, "k"] \arrow[d, "q"] & E \arrow[d, "p"] \\
C \arrow[r, "h"] & B
\end{tikzcd}$$

where p is a fibration, and a 2-cell $\alpha: f \to g$ between morphisms $f, g: A \rightrightarrows C$. By lifting the horizontal composite $h \circ \alpha$, we obtain a morphism $l(h \circ \alpha): g^*F \to f^*F$ and a 2-cell $\tilde{\alpha}: k \circ q^*f \circ l(h \circ \alpha) \to k \circ q^*g$. Using the fact that F is a **Cat**-enriched limit, it is easy to see that this 2-cell may be uniquely factored as $k \circ \beta$; and the resulting 2-cell β, together with the 1-cell $l(h \circ \alpha)$, provides the structure needed to make q into a fibration. □

A further condition which we shall wish to impose on (op)fibrations is a 'commutativity condition' between pullback along a fibration and cocomma objects, when the latter exist in our 2-category. Suppose $\mathfrak{K}$ has cocomma objects, and let $p: E \to B$ be a fibration in $\mathfrak{K}$. Given a 2-cell $\alpha: f \to g$ between two 1-cells $f, g: A \rightrightarrows B$, we now have two possible constructions. We may form the induced 1-cell $\ulcorner \alpha \urcorner: A \otimes \mathbf{2} \to B$ (recall that $A \otimes \mathbf{2}$ is just the cocomma object $(1_A \uparrow 1_A)$) and then form the pullback $(A \otimes \mathbf{2}) \times_B E$; or we may form the lifting $l(\alpha): g^*E \to f^*E$ of α and the cocomma object $(l(\alpha) \uparrow 1_{g^*E})$. It is easy to verify that the 2-cell $\tilde{\alpha}$ and the isomorphism $(p^*f)l(\alpha) \cong p^*g$ induce a canonical 1-cell

$k(\alpha)\colon (l(\alpha)\uparrow 1_{g^*E}) \to (A\otimes\mathbf{2})\times_B E$; we shall say that the fibration p satisfies condition (PCC) ('pullbacks commute with cocommas') if this comparison map is an equivalence for all α.

Lemma 4.4.4 *Let $\mathfrak{K}$ be a 2-category which has pullbacks and cocomma objects. If $p\colon E \to B$ is a fibration in $\mathfrak{K}$ satisfying condition* (PCC), *then any pullback of p also satisfies* (PCC).

Proof Let

$$\begin{tikzcd} F \arrow[r, "k"] \arrow[d, "q"] & E \arrow[d, "p"] \\ C \arrow[r, "h"] & B \end{tikzcd}$$

be such a pullback, and let $\alpha\colon f \to g$ be a 2-cell between morphisms $f, g\colon A \rightrightarrows C$. By 4.4.3(ii), we know q is a fibration; and we obtain a commutative diagram

$$\begin{tikzcd} (l(\alpha)\uparrow 1_{g^*F}) \arrow[r] \arrow[d] & F \arrow[r, "k"] \arrow[d, "q"] & E \arrow[d, "p"] \\ A\otimes\mathbf{2} \arrow[r, "\ulcorner\alpha\urcorner"] & C \arrow[r, "h"] & B \end{tikzcd}$$

where the outer rectangle is a pullback since p satisfies (PCC), and the right-hand square is a pullback by definition. So the left-hand square is a pullback; i.e. q satisfies (PCC). □

Condition (PCC) is also stable under composition of fibrations, provided the first factor of the composite is exponentiable; we shall not prove this result here, since we do not need it, but we refer the reader to [536] for the proof.

In identifying the fibrations and opfibrations in $\mathfrak{BTop}$, the following 'tensor condition' plays an important rôle. Recall from 3.4.2 that, for any topos $\mathcal{S}$, the tensor $\mathbf{2}\otimes\mathcal{S}$ in $\mathfrak{Top}$ is simply the functor category $[\mathbf{2}, \mathcal{S}]$, also called the *Sierpiński topos* over $\mathcal{S}$. We shall write $d_0\colon \mathcal{S} \to [\mathbf{2}, \mathcal{S}]$ and $d_1\colon \mathcal{S} \to [\mathbf{2}, \mathcal{S}]$ for the geometric morphisms which are respectively the source and target of the universal 2-cell with domain $\mathcal{S}$; of course, they are the geometric morphisms induced by the two internal functors $\mathbf{1} \rightrightarrows \mathbf{2}$ in $\mathcal{S}$. (Explicitly, d_j^* sends an object $(I_0 \to I_1)$ of $[\mathbf{2}, \mathcal{S}]$ to I_j; the direct image functor d_{0*} sends I to $(I \to 1)$, and $d_{1*}(I) = 1_I$.)

Definition 4.4.5 Let $\mathcal{P}$ be a class of bounded geometric morphisms which is stable under pullback, We shall say that $\mathcal{P}$ satisfies the *covariant* (resp. *contravariant*) *tensor condition* if, whenever we have a morphism $p\colon \mathcal{E} \to [\mathbf{2}, \mathcal{S}]$ in

$\mathcal{P}$ whose codomain is a Sierpiński topos, and we form the pullbacks

$$\begin{tikzcd} \mathcal{E}_0 \arrow[r, "t_0"] \arrow[d, "p_0"] & \mathcal{E} \arrow[d, "p"] & \mathcal{E}_1 \arrow[l, "p_1"'] \arrow[d, "p_1"] \\ \mathcal{S} \arrow[r, "d_0"] & {[\mathbf{2}, \mathcal{S}]} & \mathcal{S} \arrow[l, "d_1"'] \end{tikzcd} ,$$

the composite $t_0^* t_{1*}$ is the inverse (resp. direct) image of a geometric morphism over $\mathcal{S}$.

Remark 4.4.6 If the morphisms in $\mathcal{P}$ also satisfy the left (resp. right) Beck–Chevalley condition in the sense of C2.4.16 (that is, if whenever we have a pullback square

$$\begin{tikzcd} \mathcal{F} \arrow[r, "g"] \arrow[d, "q"] & \mathcal{E} \arrow[d, "p"] \\ \mathcal{T} \arrow[r, "f"] & \mathcal{S} \end{tikzcd}$$

with p, and hence q, in $\mathcal{P}$, then the Beck–Chevalley natural transformation $p^* f_* \to g_* q^*$ (resp. $f^* p_* \to q_* g^*$) is an isomorphism), then the words 'over $\mathcal{S}$' at the end of the above definition are redundant. For it is clear that the composite $d_0^* d_{1*}$ is the identity on $\mathcal{S}$, and the left (resp. right) Beck–Chevalley condition implies that we have $t_0^* t_{1*} p_1^* \cong p_0^* d_0^* d_{1*} \cong p_0^*$ (resp. $p_{1*} t_0^* t_{1*} \cong d_0^* d_{1*} p_{0*} \cong p_{0*}$). In practice, it turns out that those classes of geometric morphisms for which we are able to verify the covariant (resp. contravariant) tensor condition tend also to satisfy the appropriate-handed Beck–Chevalley condition (the classes of morphisms satisfying these conditions are characterized in C3.3.16 and C3.4.11 respectively), but there does not seem to be any necessity for this.

Theorem 4.4.7 *Let $p: \mathcal{E} \to \mathcal{S}$ be a bounded geometric morphism. Then p is a fibration (resp. an opfibration) in $\mathfrak{BTop}$ satisfying condition* (PCC) *iff the class $\mathcal{P}$ of all pullbacks of p satisfies the contravariant (resp. covariant) tensor condition.*

Proof The two cases are very similar; we shall deal with the fibration one.

First suppose the tensor condition is satisfied. Given a 2-cell $\alpha: f \to g$ between geometric morphisms $\mathcal{T} \rightrightarrows \mathcal{S}$, we form the induced morphism

$\ulcorner\alpha\urcorner\colon [\mathbf{2}, \mathcal{T}] \to \mathcal{S}$ and the pullback

$$
\begin{tikzcd}
\mathcal{F} \arrow[r, "s"] \arrow[d, "q"] & \mathcal{E} \arrow[d, "p"] \\
{[\mathbf{2}, \mathcal{T}]} \arrow[r, "\ulcorner\alpha\urcorner"] & \mathcal{S}
\end{tikzcd}
.
$$

Then the pullbacks of p along f and g are respectively the morphisms q_0 and q_1 which occur as the pullbacks of q along d_0 and d_1; so we have a geometric morphism $l(\alpha)\colon g^*\mathcal{E} \to f^*\mathcal{E}$ over $\mathcal{T}$ whose direct image is the composite $t_0^* t_{1*}$. To construct the required 2-cell $\tilde{\alpha}\colon (p^*f)l(\alpha) \to p^*g$, we observe that it suffices to construct a natural transformation $(p^*f)^* \to (l(\alpha))_*(p^*g)^*$. But p^*f and p^*g are respectively isomorphic to the composites st_0 and st_1, so this is the same thing as a natural transformation $t_0^*s^* \to t_0^*t_{1*}t_1^*s^*$; and the latter is easily obtained from the unit of $(t_1^* \dashv t_{1*})$. The verification that the constructions we have given for $l(\alpha)$ and $\tilde{\alpha}$ satisfy the compatibility conditions of Definition 4.4.1 is straightforward.

To verify condition (PCC), we observe that since d_1 and d_0 define complementary open and closed subtoposes of $[\mathbf{2}, \mathcal{T}]$, their pullbacks t_1 and t_0 define complementary open and closed subtoposes of $\mathcal{F}$; so by A4.5.6 $\mathcal{F}$ may be reconstructed by glueing along the functor $t_0^*t_{1*}$. But by 3.4.2 the cocomma object $(l(\alpha) \uparrow 1_{g^*\mathcal{E}})$ is obtained by glueing along $(l(\alpha))_*$; so it is equivalent to $\mathcal{F}$, and the construction given for $\tilde{\alpha}$ is sufficient to ensure that this equivalence is the one appearing in the statement of (PCC).

Conversely, suppose p is a fibration satisfying (PCC). Then so is any pullback of p, by 4.4.3(ii) and 4.4.4; so it suffices to verify that if the domain of p itself is of the form $[\mathbf{2}, \mathcal{T}]$ then the composite $t_0^*t_{1*}$ is a direct image functor over $\mathcal{T}$. We have a 2-cell $\alpha\colon d_0 \to d_1$ with codomain $[\mathbf{2}, \mathcal{T}]$; so $l(\alpha)$ is a geometric morphism $\mathcal{E}_1 \to \mathcal{E}_0$ over $\mathcal{T}$, and condition (PCC) tells us precisely that the topos obtained by Artin glueing along the direct image of this morphism is equivalent to $\mathcal{E}$ (which can, once again, be recovered from $\mathcal{E}_0$ and $\mathcal{E}_1$ by glueing along the composite $t_0^*t_{1*}$). Moreover, this equivalence must commute with the geometric morphisms from $\mathcal{E}_0$ and $\mathcal{E}_1$ into the two glued toposes; so it induces an isomorphism $l(\alpha)_* \cong t_0^*t_{1*}$. □

One advantage of 4.4.7 is that it enables us to prove that (op)fibrations in $\mathfrak{BTop}$ are stable under adjoint or coadjoint retracts, in the sense of 1.1.9(c).

Lemma 4.4.8 *Let $\mathcal{P}$ be a class of bounded geometric morphisms satisfying either the covariant or the contravariant tensor condition. Let $\mathcal{Q}$ be either the class of morphisms defined by saying that $f\colon \mathcal{E} \to \mathcal{S}$ is in $\mathcal{Q}$ iff it is an adjoint retract in $\mathfrak{BTop}/\mathcal{S}$ of a morphism in $\mathcal{P}$, or the class of coadjoint retracts of morphisms in $\mathcal{P}$. Then $\mathcal{Q}$ satisfies the same tensor condition as $\mathcal{P}$.*

Proof There are four different cases to consider, though two of them are formally dual to the other two. However, all four proofs are very similar, and we shall give only one in detail, namely the case when $\mathcal{P}$ satisfies the covariant tensor condition and we consider adjoint retracts. (The case of coadjoint retracts and the covariant tensor condition is treated in [536].)

We note first that $\mathcal{Q}$ is stable under pullback, since pullback along a fixed morphism is a 2-functor and hence preserves adjoint retracts. So suppose $p\colon \mathcal{E} \to [\mathbf{2},\mathcal{S}]$ is a morphism in $\mathcal{P}$, and let $q\colon \mathcal{F} \to [\mathbf{2},\mathcal{S}]$ be a retract of it by morphisms $r\colon \mathcal{E} \to \mathcal{F}$ and $i\colon \mathcal{F} \to \mathcal{E}$ over $[\mathbf{2},\mathcal{S}]$ such that r is left adjoint to i in $\mathfrak{Top}$, i.e. such that $r^* \cong i_*$. Form the diagram

$$\begin{tikzcd}
\mathcal{F}_0 \arrow[r, "u_0"] \arrow[d, "i_0"] & \mathcal{F} \arrow[d, "i"] & \mathcal{F}_1 \arrow[l, "u_1"'] \arrow[d, "i_1"] \\
\mathcal{E}_0 \arrow[r, "t_0"] \arrow[d, "r_0"] & \mathcal{E} \arrow[d, "r"] & \mathcal{E}_1 \arrow[l, "t_1"'] \arrow[d, "r_1"] \\
\mathcal{F}_0 \arrow[r, "u_0"] \arrow[d, "q_0"] & \mathcal{F} \arrow[d, "q"] & \mathcal{F}_1 \arrow[l, "u_1"'] \arrow[d, "q_1"] \\
\mathcal{S} \arrow[r, "d_0"] & {[\mathbf{2},\mathcal{S}]} & \mathcal{S} \arrow[l, "d_1"']
\end{tikzcd}$$

where all the squares are pullbacks: then we know that the composite $t_0^* t_{1*}$ is an inverse image functor over $\mathcal{S}$, and we also have $(r_0 \dashv i_0)$ and $(r_1 \dashv i_1)$. Now

$$\begin{aligned}
u_0^* u_{1*} &\cong i_0^* i_{0*} u_0^* u_{1*} \\
&\cong i_0^* r_0^* u_0^* u_{1*} \\
&\cong i_0^* t_0^* r^* u_{1*} \\
&\cong i_0^* t_0^* i_* u_{1*} \\
&\cong i_0^* t_0^* t_{1*} i_{i*} \\
&\cong i_0^* t_0^* t_{1*} r_1^*,
\end{aligned}$$

which is a composite of inverse image functors over $\mathcal{S}$. □

We are now ready to consider two particular classes of morphisms in $\mathfrak{BTop}$ for which we can verify the tensor condition. The first is the class of morphisms

of the form $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$:

Lemma 4.4.9 *For any internal category $\mathbb{C}$ in a topos $\mathcal{S}$, the canonical morphism $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$ is an opfibration in $\mathfrak{BTop}$ satisfying condition* (PCC).

Proof We verify that the class of all such morphisms satisfies the covariant tensor condition. By 3.2.14, we know it is stable under pullback. Also, an internal category $\mathbb{C}$ in $[\mathbf{2}, \mathcal{S}]$ is the same thing as an internal functor $f: \mathbb{C}_0 \to \mathbb{C}_1$ in $\mathcal{S}$, and an easy calculation shows that, in the diagram

$$\begin{tikzcd}
{[\mathbb{C}_0, \mathcal{S}]} \arrow[r, "t_0"] \arrow[d] & {[\mathbb{C}, [\mathbf{2}, \mathcal{S}]]} \arrow[d] & {[\mathbb{C}_1, \mathcal{S}]} \arrow[l, "t_1"'] \arrow[d] \\
\mathcal{S} \arrow[r, "d_0"] & {[\mathbf{2}, \mathcal{S}]} & \mathcal{S} \arrow[l, "d_1"']
\end{tikzcd} ,$$

the composite $t_0^* t_{1*}$ is simply the functor obtained by applying pullback along f to discrete opfibrations over $\mathbb{C}_1$, and hence it is the inverse image of a geometric morphism over $\mathcal{S}$. □

The toposes which occur as retracts in $\mathfrak{BTop}/\mathcal{S}$ of diagram categories $[\mathbb{C}, \mathcal{S}]$ are studied in Section C4.3, where we shall in particular prove that every such $\mathcal{S}$-topos is in fact a coadjoint retract of a diagram category (cf. C4.3.9). Thus we may conclude:

Corollary 4.4.10 *Let $p: \mathcal{E} \to \mathcal{S}$ be a bounded geometric morphism which is a retract in $\mathfrak{BTop}/\mathcal{S}$ of one of the form $[\mathbb{C}, \mathcal{S}] \to \mathcal{S}$. Then p is an opfibration in $\mathfrak{BTop}$ satisfying condition* (PCC). □

Similarly, it turns out that coherent morphisms, as considered at the end of the previous section, are fibrations in $\mathfrak{Top}/\mathcal{S}$. Once again, we shall have to rely on a couple of results from Part C in order to prove this.

Lemma 4.4.11 *Let $p: \mathcal{E} \to \mathcal{S}$ be a coherent topos over $\mathcal{S}$. Then p is a fibration in $\mathfrak{BTop}$ satisfying condition* (PCC).

Proof We noted in 4.3.9 that the class of coherent morphisms is stable under pullback, so we have only to verify the tensor condition for a coherent morphism $\mathcal{E} \simeq \mathbf{Sh}_{[\mathbf{2}, \mathcal{S}]}(\mathbb{C}, T) \to [\mathbf{2}, \mathcal{S}]$. Consider first the case when the coverage T is trivial, so that $\mathcal{E}$ is simply the diagram category $[\mathbb{C}^{\mathrm{op}}, [\mathbf{2}, \mathcal{S}]]$: in this case, as in 4.4.9, we find that $t_0^* t_{1*}$ is simply the functor obtained by pulling back discrete fibrations along $f: \mathbb{C}_0 \to \mathbb{C}_1$. But $\mathbb{C}$ is now assumed to be a cartesian internal category in $[\mathbf{2}, \mathcal{S}]$, which means that $\mathbb{C}_0$ and $\mathbb{C}_1$ are cartesian categories in $\mathcal{S}$ and f is a cartesian functor; so, by the $\mathcal{S}$-indexed version of A4.1.10 (cf. also 3.2.8(c)), pullback along f is the direct image of a geometric morphism over $\mathcal{S}$.

For the general case, we have to presume upon the reader's goodwill by appealing to results in Part C: this time in Section C2.3, where we develop a

notion of 'cover-preserving functor' and show that if $f: (\mathbb{C}_0, T_0) \to (\mathbb{C}_1, T_1)$ is a cartesian cover-preserving functor between cartesian sites, then the geometric morphism $[\mathbb{C}_0^{\mathrm{op}}, \mathcal{S}] \to [\mathbb{C}_1^{\mathrm{op}}, \mathcal{S}]$ whose direct image is given by composition with f restricts to a morphism between the corresponding sheaf toposes (see C2.3.4). It is clear that this result applies in the present case: the assertion that T is an internal coverage in $[\mathbf{2}, \mathcal{S}]$ contains the information that f preserves covers. □

Once again, 4.4.11 may be extended to the class of all morphisms $p: \mathcal{E} \to \mathcal{S}$ such that $\mathcal{E}$ is either an adjoint or a coadjoint retract in $\mathfrak{BTop}/\mathcal{S}$ of a coherent $\mathcal{S}$-topos. We do not know whether this class includes all retracts of coherent $\mathcal{S}$-toposes; but for the particular case of localic morphisms, we study those which are retracts of coherent ones in Section C4.1, and in C4.1.13 we show that they all occur as coadjoint retracts. (We characterize them there as the class of morphisms corresponding to *stably locally compact* locales.) Thus we may state

Corollary 4.4.12 *If X is a stably locally compact locale in a topos $\mathcal{S}$, then the canonical geometric morphism* $\mathbf{Sh}_{\mathcal{S}}(X) \to \mathcal{S}$ *is a fibration in* $\mathfrak{BTop}$ *satisfying condition* (PCC). □

The other notion we wish to introduce in this section is a 2-dimensional generalization of the notion of partial product, which we studied in A1.5.7. It is clear that, for an arbitrary morphism $f: E \to B$ in a 2-category $\mathfrak{K}$, and an object A of $\mathfrak{K}$, it does not make much sense to ask whether the partial product $P(A, f)$ exists, since we cannot hope to make it universal in a 2-dimensional sense. That is, we wish 1-cells $C \to P(A, f)$ to correspond to pairs $(C \to B, C \times_B E \to A)$ of 1-cells, but we cannot say what 2-cells between such 1-cells should correspond to, since a 2-cell between morphisms $C \rightrightarrows P(A, f)$ would induce a 2-cell between morphisms $C \rightrightarrows B$ by horizontal composition, but we have no means of pulling the latter back along f. However, if f has either a fibration or an opfibration structure, in the sense defined above, we can circumvent this problem.

For definiteness, let us suppose that f is an opfibration. Then, for any object C of $\mathfrak{K}$, we define a category of 'partial product cones over (A, f) with vertex C', as follows: its objects are pairs (u, v) with $u: C \to B$ and $v: C \times_B E \to A$, and a morphism $(u, v) \to (u', v')$ is a pair (α, β) where α is a 2-cell $u \to u'$ and β is a 2-cell $v \to v' \cdot l(\alpha)$. The assignment mapping C to this category is (contravariantly) pseudo-functorial in C, in an obvious way; and by a *covariant partial product* $\mathcal{P}_\bullet(A, f)$, we mean a representing object for this functor. Similarly, if f is a fibration, we may define the notion of *contravariant partial product*, for which our notation is $\mathcal{P}^\bullet(A, f)$.

The next result is the '2-dimensional analogue' of the basic existence theorem A1.5.7 for partial products.

Theorem 4.4.13 *Let $\mathfrak{K}$ be a 2-category with pullbacks and cocomma objects, and suppose that $p: E \to B$ is a fibration (resp. an opfibration) in $\mathfrak{K}$ which satisfies condition* (PCC) *and is exponentiable as an object of $\mathfrak{K}/B$. Suppose also*

that products with B exist in $\mathfrak{K}$. Then, for any object A of $\mathfrak{K}$, the contravariant (resp. covariant) partial product $\mathcal{P}^\bullet(A,p)$ (resp. $\mathcal{P}_\bullet(A,p)$) exists in $\mathfrak{K}$. Moreover, $\mathcal{P}^\bullet(-,p)$ (resp. $\mathcal{P}_\bullet(-,p)$) is covariantly functorial, and may be described as the composite

$$\mathfrak{K} \xrightarrow{B^*} \mathfrak{K}/B \xrightarrow{(-)^p} \mathfrak{K}/B \xrightarrow{\Sigma_B} \mathfrak{K},$$

where Σ_B and B^* as usual denote the forgetful functor $\mathfrak{K}/B \to \mathfrak{K}$ and its right adjoint. In particular, if p is both a fibration and an opfibration, and satisfies (PCC) *in both senses), then the two notions of partial product coincide.*

Proof We shall consider the opfibration case; that for fibrations is similar. Given an object C of $\mathfrak{K}$, we have to construct a natural equivalence between the category of partial product cones over (A,p) with vertex C, and the category $\mathfrak{K}(C, \Sigma_B((B^*A)^p))$. The correspondence on (isomorphism classes of) objects is easy to describe: specifying a morphism $f\colon C \to \Sigma_B((B^*A)^p)$ in $\mathfrak{K}$ is equivalent to specifying a morphism $g\colon C \to B$ in $\mathfrak{K}$ and a morphism $g \to (B^*A)^p$, or equivalently $g \times p \to B^*A$, in $\mathfrak{K}/B$. And the latter further corresponds to a morphism $\Sigma_B(g \times p) = C \times_B E \to A$ in $\mathfrak{K}$. To show that this correspondence extends to an equivalence of categories, we use (PCC) to show that a morphism between two partial product cones with vertex C may be regarded as a single partial product cone with vertex $C \otimes \mathbf{2}$; the details are straightforward. The final assertion of the theorem follows from the observation that the displayed composite takes no account of the variance of p. □

In the previous section, we verified the exponentiability in $\mathfrak{BTop}/\mathcal{E}$ of all the morphisms $\mathcal{F} \to \mathcal{E}$ which we have proved to be (op)fibrations in the present section. So we may now conclude

Corollary 4.4.14 *Let $\mathcal{S}$ be a topos with a natural number object, $f\colon \mathcal{F} \to \mathcal{E}$ a morphism of $\mathfrak{BTop}/\mathcal{S}$ and $\mathcal{G}$ a bounded $\mathcal{S}$-topos. Then*

(i) *If f is a retract in $\mathfrak{BTop}/\mathcal{E}$ of a topos of the form $[\mathbb{C}, \mathcal{E}]$ for some internal category $\mathbb{C}$ in $\mathcal{E}$, then the covariant partial product $\mathcal{P}_\bullet(\mathcal{G}, f)$ exists in $\mathfrak{B}\mathcal{T}op/\mathcal{S}$.*

(ii) *If f is either an adjoint or a coadjoint retract of a coherent $\mathcal{E}$-topos, then the contravariant partial product $\mathcal{P}^\bullet(\mathcal{G}, f)$ exists in $\mathfrak{BTop}/\mathcal{S}$.*

(iii) *In particular, if f is a retract of an $\mathcal{E}$-topos of the form $[\mathbb{C}^{\mathrm{op}}, \mathcal{E}]$ where $\mathbb{C}$ is cartesian, then the two partial products exist and are equivalent.* □

The $\mathcal{E}$-toposes satisfying the condition of 4.4.14(iii) are characterized in C4.3.1 as the injective objects (with respect to inclusions) in the 2-category $\mathfrak{BTop}/\mathcal{E}$.

The most important example of a partial product functor is the (lower) bagdomain functor B_L, first introduced in [535] and [1206]. To explain the thinking behind this notion, suppose $\mathcal{E}$ is the classifying topos (over $\mathcal{S}$) for a geometric

theory $\mathbb{T}$. If I is a particular object of $\mathcal{S}$, we know how to construct a classifying topos for the theory of I-indexed families of $\mathbb{T}$-models, and we have seen in the previous section that this is the exponential $\mathcal{E}^{\mathcal{S}/I}$ in $\mathfrak{BTop}/\mathcal{S}$. But what if we wish to classify families of $\mathbb{T}$-models indexed by a 'variable' object?

By a *bag* (or indexed family) of $\mathbb{T}$-models in an $\mathcal{S}$-topos $\mathcal{F}$, we mean a $\mathbb{T}$-model in $\mathcal{F}/B$ for some B; if M and M' are bags of $\mathbb{T}$-models, living in the slice categories $\mathcal{F}/B$ and $\mathcal{F}/B'$, then a morphism of bags of $\mathbb{T}$-models from (B, M) to (B', M') consists of a morphism $f\colon B \to B'$ in $\mathcal{F}$ together with a morphism $M \to f^*M'$ in $\mathbb{T}(\mathcal{F}/B)$. It is easy to see that such morphisms can be composed, so that we have a category $\mathbb{T}\text{-bag}(\mathcal{F})$ of bags of $\mathbb{T}$-models in $\mathcal{F}$; also, a geometric morphism $g\colon \mathcal{G} \to \mathcal{F}$ over $\mathcal{S}$ induces a functor $g^*: \mathbb{T}\text{-bag}(\mathcal{F}) \to \mathbb{T}\text{-bag}(\mathcal{G})$ in the obvious way, and geometric transformations induce natural transformations between such functors. Thus $\mathbb{T}$-bag is itself a theory over $\mathcal{S}$.

In passing, we note

Lemma 4.4.15 *For any geometric theory $\mathbb{T}$ over $\mathcal{S}$, and any bounded $\mathcal{S}$-topos $\mathcal{F}$, $\mathbb{T}\text{-bag}(\mathcal{F})$ is equivalent to the category of pre-geometric morphisms $\mathcal{F} \to \mathcal{S}[\mathbb{T}]$ (in the sense of* A4.1.13*) over $\mathcal{S}$, and pre-geometric transformations between them (that is, natural transformations between their inverse image functors).*

Proof This is immediate from the canonical factorization

$$\mathcal{F} \longrightarrow \mathcal{F}/f^*(1) \xrightarrow{\hat{f}} \mathcal{S}[\mathbb{T}]$$

of an arbitrary pre-geometric morphism f, constructed in A4.1.13. (Note, however, that the statement 'f is a pre-geometric morphism over $\mathcal{S}$' simply means that its direct and inverse image functors form an $\mathcal{S}$-indexed adjoint pair; it does *not* imply that the diagram

$$\begin{tikzcd} \mathcal{F} \arrow[r, "f"] \arrow[dr] & \mathcal{S}[\mathbb{T}] \arrow[d] \\ & \mathcal{S} \end{tikzcd}$$

commutes.) □

Proposition 4.4.16 *Let $\mathcal{S}$ be a topos with a natural number object. For any geometric theory $\mathbb{T}$ over $\mathcal{S}$, the theory $\mathbb{T}$-bag is also geometric.*

Proof Given a classifying topos $\mathcal{E} = \mathcal{S}[\mathbb{T}]$ for $\mathbb{T}$, we claim that the covariant partial product $\mathcal{P}_\bullet(\mathcal{E}, f)$ has the universal property of a classifying topos for $\mathbb{T}$-bag, where f is the geometric morphism $\mathcal{S}[\mathbb{O}]/G \to \mathcal{S}[\mathbb{O}]$ and G is the generic object in $\mathcal{S}[\mathbb{O}]$. But this is immediate from the definition of a covariant partial

product; for a geometric morphism $\mathcal{F} \to \mathcal{P}_\bullet(\mathcal{E}, f)$ corresponds by definition to a pair (u, v), where u is a morphism $\mathcal{F} \to \mathcal{S}[\mathbb{O}]$ (equivalently, an object $B = u^*(G)$ of $\mathcal{F}$), and v is a morphism from the pullback $\mathcal{F} \times_{\mathcal{S}[\mathbb{O}]} \mathcal{S}[\mathbb{O}]/G \simeq \mathcal{F}/B$ to $\mathcal{E}$ (equivalently, a $\mathbb{T}$-model in $\mathcal{F}/B$). A similarly straightforward argument shows that a geometric transformation between two morphisms $\mathcal{F} \rightrightarrows \mathcal{P}_\bullet(\mathcal{E}, f)$ corresponds to a morphism of bags of $\mathbb{T}$-models, as defined above. $\square$

We write $B_L\mathcal{E}$ for the topos $\mathcal{P}_\bullet(\mathcal{E}, f)$ introduced in the proof of 4.4.16, and call it the *lower* (or covariant) *bagdomain* of $\mathcal{E}$. Clearly, the construction $\mathcal{E} \mapsto B_L\mathcal{E}$ is functorial on $\mathfrak{BTop}/\mathcal{S}$.

Example 4.4.17 If $\mathcal{E}$ is $\mathcal{S}$ itself, then $B_L\mathcal{E}$ is simply the object classifier $\mathcal{S}[\mathbb{O}]$. More generally, suppose $\mathcal{E}$ is $[\mathbb{C}, \mathcal{S}]$ for some internal category $\mathbb{C}$ in $\mathcal{S}$; then we claim that $B_L\mathcal{E}$ may be identified with $[\mathrm{Fam}_f(\mathbb{C}), \mathcal{S}]$ where $\mathrm{Fam}_f(\mathbb{C})$ is the free finite-coproduct completion of $\mathbb{C}$, introduced in 1.4.17. (Recall that $\mathbb{S}_f \cong \mathrm{Fam}_f(\mathbf{1})$, so this is indeed a generalization of the case $\mathcal{E} = \mathcal{S}$.) To prove this claim, note that a bag of $\mathbb{C}$-torsors in $\mathcal{F}$ (that is, a $\mathbb{C}$-torsor in some slice category of $\mathcal{F}$) may by 2.6.6 be identified with a discrete fibration $\mathbb{F} \to \mathbb{C}$ in $\mathcal{F}$ whose domain is a weakly filtered category. We shall show that, if this holds, then the induced functor $\mathrm{Fam}_f(\mathbb{F}) \to \mathrm{Fam}_f(\mathbb{C})$ is a $\mathrm{Fam}_f(\mathbb{C})$-torsor.

For this, we argue in terms of elements, as if $\mathcal{F}$ were the topos of sets. The objects of $\mathrm{Fam}_f(\mathbb{F})$ are finite indexed families $g\colon [p] \to F_0$ of objects of $\mathbb{F}$; clearly, there exists such a family, namely the unique morphism $[o] \cong 0 \to F_0$, and given two families g_1 and g_2, the induced map

$$[p_1 + p_2] \cong [p_1] \amalg [p_2] \xrightarrow{(g_1, g_2)} F_0$$

is an object of $\mathrm{Fam}_f(\mathbb{F})$ to which each of them can be mapped. Given two objects g_1 and g_2 of $\mathrm{Fam}_f(\mathbb{F})$, a parallel pair of morphisms $\alpha, \beta\colon g_1 \rightrightarrows g_2$ yields in particular a parallel pair of morphisms $[p_1] \rightrightarrows [p_2]$ between the indexing cardinals; if we write $[p_2] \twoheadrightarrow [p_3]$ for the coequalizer of the latter (cf. D5.2.5), then for each element of $[p_3]$ we have a *connected* finite diagram in $\mathbb{F}$ consisting of the objects $g_1(i)$ and $g_2(j)$ for which i and j map to the given element of $[p_3]$. If we choose a cone under this diagram for each $k : [p_3]$, we obtain morphism $g_3\colon [p_3] \to F_0$ (picking out the vertices of these cones) for which we have a morphism $\gamma\colon g_2 \to g_3$ in $\mathrm{Fam}_f(\mathbb{F})$ coequalizing α and β. So $\mathrm{Fam}_f(\mathbb{F})$ is filtered.

Conversely, if we are given a $\mathrm{Fam}_f(\mathbb{C})$-torsor $\mathbb{G}$, then the pullback of $\mathbb{G}$ along the inclusion $\mathbb{C} \to \mathrm{Fam}_f(\mathbb{C})$ (obtained by regarding objects as $[so]$-indexed families) is easily seen to be a weakly filtered category $\mathbb{F}$; and we have an isomorphism $\mathbb{G} \cong \mathrm{Fam}_f(\mathbb{F})$, since the functor corresponding to $\mathbb{G}$ must send coproducts in $\mathrm{Fam}_f(\mathbb{C})$ to products. Thus we have established an equivalence between the category of $\mathrm{Fam}_f(\mathbb{C})$-torsors and the category of 'pre-torsors' (that is, discrete fibrations with weakly filtered domain) over $\mathbb{C}$.

Generalizing Example 4.4.17, it is possible to proceed from a site of definition $(\mathbb{C}, J)$ for an arbitrary bounded $\mathcal{S}$-topos $\mathcal{E}$, to one for $B_L\mathcal{E}$ whose underlying category is $(\mathrm{Fam}_f(\mathbb{C}^{\mathrm{op}}))^{\mathrm{op}}$. We omit the details, which may be found in [535].

Remark 4.4.18 Lemma 4.4.15 may now be reformulated as saying that $B_L\mathcal{E}$ is the 'free coproduct completion' of $\mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$, in the sense that, for any $\mathcal{S}$-topos $\mathcal{F}$, the category $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F}, B_L\mathcal{E})$ of '$\mathcal{F}$-valued points of $B_L\mathcal{E}$' is the free coproduct completion of $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F}, \mathcal{E})$ as an $\mathcal{F}$-indexed category (cf. 1.4.16).

In view of this observation, it will come as no surprise to discover that the functor B_L has the structure of a KZ-monad on $\mathfrak{BTop}/\mathcal{S}$. The unit map $\eta_\mathcal{E}\colon \mathcal{E} \to B_L\mathcal{E}$ is the morphism which arises from regarding a $\mathbb{T}$-model as a bag of $\mathbb{T}$-models indexed by a singleton: explicitly, it corresponds to the pair of geometric morphisms $(u, 1_\mathcal{E})$ where $u\colon \mathcal{E} \to \mathcal{S}[\mathbb{O}]$ classifies the terminal object of $\mathcal{E}$. (Equivalently again, in terms of 4.4.15, it corresponds to the identity geometric morphism $\mathcal{E} \to \mathcal{E}$ regarded as a pre-geometric morphism.) The multiplication similarly corresponds to the operation which sends a bag of bags of $\mathbb{T}$-models $((M_i \mid i \in I_j) \mid j \in J)$ to the single bag $(M_i \mid i \in \coprod_{j\in J} I_j)$; equivalently, it corresponds to the pre-geometric morphism $B_L(B_L\mathcal{E}) \to \mathcal{E}$ displayed as the top edge and vertical composite in the diagram

$$\begin{tikzcd}
 & B_L(B_L\mathcal{E})/C \arrow[r] \arrow[d] & B_L\mathcal{E}/B_1 \arrow[r] \arrow[d] & \mathcal{E} \\
\mathcal{S}[\mathbb{O}]/G \arrow[d] & B_L(B_L\mathcal{E})/B_2 \arrow[l] \arrow[r, "v_2"] \arrow[d] & B_L\mathcal{E} & \\
\mathcal{S}[\mathbb{O}] & B_L(B_L\mathcal{E}) \arrow[l, "u_2"'] & &
\end{tikzcd}$$

where (u_1, v_1) and (u_2, v_2) denote the structure morphisms associated with the partial products $B_L\mathcal{E}$ and $B_L(B_L\mathcal{E})$ respectively, $B_i = u_i^*(G)$ and $(C \to B_2) = v_2^*(B_1)$. The verification that this is indeed a KZ-monad structure on B_L is straightforward.

We may characterize the algebras for this monad in several different ways:

Theorem 4.4.19 *Let $\mathcal{S}$ be a topos with a natural number object, and $p\colon \mathcal{E} \to \mathcal{S}$ a bounded $\mathcal{S}$-topos. The following are equivalent:*

(i) *$\mathcal{E}$ admits an algebra structure for the bagdomain monad B_L.*

(ii) *$\eta_\mathcal{E}\colon \mathcal{E} \to B_L\mathcal{E}$ has a left adjoint in $\mathfrak{BTop}/\mathcal{S}$.*

(iii) *$\eta_\mathcal{E}\colon \mathcal{E} \to B_L\mathcal{E}$ admits a splitting in $\mathfrak{BTop}/\mathcal{S}$.*

(iv) *$\mathcal{E}$ is a retract in $\mathfrak{BTop}/\mathcal{S}$ of a topos of the form $B_L\mathcal{F}$.*

(v) *For each object I of $\mathcal{S}$, the 'diagonal' morphism $\mathcal{E} \to \mathcal{E}^{\mathcal{S}/I}$ (that is, the exponential transpose of the projection $\mathcal{E} \times_{\mathcal{S}} \mathcal{S}/I \simeq \mathcal{E}/p^*I \to \mathcal{E}$) has a left adjoint in $\mathfrak{BTop}/\mathcal{S}$.*

(vi) *The morphism p and the diagonal $\Delta_p \colon \mathcal{E} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$ both have left adjoints in $\mathfrak{BTop}/\mathcal{S}$.*

(vii) *For each bounded $\mathcal{S}$-topos $q \colon \mathcal{F} \to \mathcal{S}$, the category $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ has finite coproducts, which are natural in $\mathcal{F}$.*

(viii) *For each bounded $\mathcal{S}$-topos $q \colon \mathcal{F} \to \mathcal{S}$, the inclusion of $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ in the category $\mathfrak{PreGeom}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ of pre-geometric morphisms over $\mathcal{S}$ from $\mathcal{F}$ to $\mathcal{E}$ (cf.* 4.4.15*) has a left adjoint, and these adjoints are natural in $\mathcal{F}$.*

Proof Suppose $\mathcal{E}$ is the classifying topos for a geometric theory $\mathbb{T}$ over $\mathcal{S}$. It is easy to see that $\eta_{\mathcal{E}}$ is always an inclusion, since (in terms of classifying toposes) it corresponds to forcing a bag of $\mathbb{T}$-models to satisfy the geometric axiom (in the sense of 4.2.7(b)) which says that the object indexing the bag is a singleton. Hence, if it has a left adjoint, the latter is necessarily a one-sided inverse for it; thus the equivalence of (i), (ii) and (iii) follows from the general theory of KZ-monads (1.1.13 and 1.1.14), plus the fact that $\mathfrak{BTop}/\mathcal{S}$ is locally Cauchy-complete (cf. A4.1.15). (iii) $\Rightarrow$ (iv) is immediate, and (iv) $\Rightarrow$ (ii) follows from 1.1.10.

(ii) $\Rightarrow$ (v): Clearly, the exponential $\mathcal{E}^{\mathcal{S}/I}$ classifies the theory of I-indexed families of $\mathbb{T}$-models: thus we have a canonical geometric morphism $\mathcal{E}^{\mathcal{S}/I} \to B_L\mathcal{E}$, which can be described in terms of classifying toposes as the effect of making the simple $\mathcal{S}$-specialization (4.2.8(d)) of $\mathbb{T}$-bag which forces the object indexing the generic bag of $\mathbb{T}$-models to be isomorphic to the constant geometric construct I. It is easy to see (by considering the effect on families of $\mathbb{T}$-models) that if we compose this morphism with the left adjoint $B_L\mathcal{E} \to \mathcal{E}$ of $\eta_{\mathcal{E}}$, we obtain a left adjoint for the diagonal.

(v) $\Rightarrow$ (vi) is immediate, since (vi) is equivalent to the two special cases $I = 0$ and $I = 1 \amalg 1$ of (v). And (vi) $\Rightarrow$ (vii) is also easy, since (for example) we may form the coproduct in $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ of a pair of morphisms $\mathcal{F} \rightrightarrows \mathcal{E}$ by regarding them as a single morphism $\mathcal{F} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$ and then composing with the left adjoint of the diagonal.

For (vii) $\Rightarrow$ (viii), we note that (viii) is essentially the statement that $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ has '$\mathcal{F}$-indexed coproducts which are natural in $\mathcal{F}$', where we make it into an $\mathcal{F}$-indexed category by defining its B-indexed families of objects (for an object B of $\mathcal{F}$) to be geometric morphisms $\mathcal{F}/B \to \mathcal{E}$ over $\mathcal{S}$ (equivalently, by identifying it with the category $\mathfrak{BTop}/\mathcal{F}\,(\mathcal{F},\mathcal{F} \times_{\mathcal{S}} \mathcal{E})$, and indexing the latter over $\mathcal{F}$ in the manner described in Section B3.4). But by 3.4.8 this category always has colimits indexed by filtered internal categories in $\mathcal{F}$; so the result we need is just an $\mathcal{F}$-indexed version of the fact that arbitrary coproducts may be constructed from finite coproducts and filtered colimits (cf. D5.2.14).

Finally, (viii) implies (ii) by applying the left adjoint of (viii) to the universal pre-geometric morphism from $B_L\mathcal{E}$ to $\mathcal{E}$. $\square$

An $\mathcal{S}$-topos satisfying the equivalent conditions of 4.4.19 is called *hyperlocal*. (The reason for this name will become clear in Section C3.6, when we consider the class of local morphisms – they are the morphisms $p: \mathcal{E} \to \mathcal{S}$ for which p has a left adjoint in $\mathfrak{Top}/\mathcal{S}$, although Δ_p may fail to have a left adjoint. We shall see in C3.6.5 that bounded local $\mathcal{S}$-toposes can also be characterized as the algebras for a KZ-monad on $\mathfrak{BTop}/\mathcal{S}$, whose functor part is a partial product functor.)

Examples 4.4.20 (*a*) If $\mathbb{C}$ is an internal category in $\mathcal{S}$ with finite coproducts, then $[\mathbb{C}, \mathcal{S}]$ is a hyperlocal $\mathcal{S}$-topos: in view of 3.2.13, the binary coproduct functor $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and the initial object of $\mathbb{C}$ induce the left adjoints required for 4.4.19(vi). Conversely if $\mathbb{C}$ is Cauchy-complete and $[\mathbb{C}, \mathcal{S}]$ is hyperlocal, then since the left adjoints of 4.4.19(vi) are automatically $\mathcal{S}$-essential it follows from the $\mathcal{S}$-indexed version of A4.1.5 that they must be induced by internal functors $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and $\mathbf{1} \to \mathbb{C}$, and hence that $\mathbb{C}$ has finite coproducts.

(*b*) To give examples of spatial hyperlocal toposes, we need to borrow some ideas from Chapter C1. Let X be a sober space whose points form a join-semilattice in the specialization ordering (defined by $x \leq y$ iff x belongs to the closure of $\{y\}$, cf. the remarks before C1.4.5), and such that the binary join map $X \times X \to X$ is continuous. (Here $X \times X$ needs to be interpreted as the product in the category **Loc** of locales rather than in **Sp**; but the two coincide if X is locally compact – cf. C4.1.8.) Then it follows easily from C1.4.5 and C1.4.8 that the geometric morphisms $\mathbf{Sh}(X \times X) \to \mathbf{Sh}(X)$ and $\mathbf{Set} \to \mathbf{Sh}(X)$ induced (as in A4.1.11) by the join map and the least element of X provide the left adjoints required for 4.4.19(vi); so $\mathbf{Sh}(X)$ is a hyperlocal **Set**-topos. In particular, this holds if X is a continuous lattice equipped with its Scott topology (cf. C4.1.6).

The use of the name '*lower* bagdomain' will naturally have aroused in the reader an expectation that we will also describe an 'upper' or contravariant bag-domain functor, and prove corresponding results about it. Such a construction indeed exists, but its properties are less well explored; we shall merely sketch the construction here, and refer the interested reader to [537] for further information.

The basic idea underlying the construction is that, just as an object of a topos $\mathcal{S}$ is canonically a filtered colimit of finite cardinals (an idea which we used in the proof of 4.4.19(vii) $\Rightarrow$ (viii)), so the objects indexing 'co-bags' of $\mathbb{T}$-models should be canonically expressible as cofiltered limits of finite cardinals. Consideration of the latter leads us naturally into the category of Stone spaces (or better, Stone locales) – cf. [520], VI 3.1. So a co-bag of $\mathbb{T}$-models in $\mathcal{S}$ should be taken to be a sheaf of $\mathbb{T}$-models on a Stone locale X in $\mathcal{S}$, i.e. a $\mathbb{T}$-model in $\mathbf{Sh}_{\mathcal{S}}(X)$, or equivalently a $\mathbb{T}$-model in $\mathbf{Sh}_{\mathcal{S}}(B, C)$, where B is an internal Boolean algebra in $\mathcal{S}$ and C is its coherent coverage (cf. A2.1.11(*b*)).

Definition 4.4.21 (*a*) We call a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ *entire* if there exists a Boolean algebra B in $\mathcal{S}$ such that $\mathcal{F}$ is equivalent (as an $\mathcal{S}$-topos) to $\mathbf{Sh}_{\mathcal{S}}(B, C)$. By 4.3.8, any entire morphism is coherent; so by 4.4.14(ii) the contravariant partial product functor $\mathcal{P}^{\bullet}(-, f)$ exists in $\mathfrak{BTop}/\mathcal{S}$, for any $\mathcal{S}$ such that $\mathcal{E}$ is bounded over $\mathcal{S}$.

(*b*) Let $\mathbb{T}$ be a geometric theory over $\mathcal{S}$, with classifying topos $p: \mathcal{E} \to \mathcal{S}$. By a *co-bag of* $\mathbb{T}$*-models* in a bounded $\mathcal{S}$-topos $\mathcal{F}$, we mean a span

$$\mathcal{F} \xleftarrow{\ g\ } \mathcal{G} \xrightarrow{\ f\ } \mathcal{E}$$

in $\mathfrak{BTop}/\mathcal{S}$, where g is entire; a morphism of co-bags of $\mathbb{T}$-models from (f, g) to (f', g') is a diagram of the form

$$\begin{tikzcd}
 & \mathcal{G} \arrow[dl, "g"'] \arrow[dr, "f"] & \\
\mathcal{F} & & \mathcal{E} \\
 & \mathcal{G}' \arrow[ul, "g'"] \arrow[ur, "f'"'] \arrow[uu] \arrow[ur, phantom, shift left=6, "\downarrow"] &
\end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$, where the left-hand cell commutes up to isomorphism but the right-hand one contains a not-necessarily-invertible 2-cell. We write $\mathbb{T}$-cobag$(\mathcal{F})$ for the category of co-bags of $\mathbb{T}$-models in $\mathcal{F}$; since pullbacks of entire morphisms are entire by 4.3.9, it is easy to see that $\mathbb{T}$-cobag is again a theory over $\mathcal{S}$.

(*c*) Let $\mathbb{B}$ denote the (algebraic) theory of Boolean algebras, and let $G_{\mathbb{B}}$ denote the generic Boolean algebra in the Boolean algebra classifier $\mathcal{B} = \mathcal{S}[\mathbb{B}]$ over $\mathcal{S}$. We define the *upper bagdomain* $B_U\mathcal{E}$ of a bounded $\mathcal{S}$-topos $\mathcal{E}$ to be $\mathcal{P}^{\bullet}(\mathcal{E}, f)$, where f is the canonical entire morphism $\mathbf{Sh}_{\mathcal{B}}(G_{\mathbb{B}}, C) \to \mathcal{B}$.

Theorem 4.4.22

(i) *If $\mathcal{E}$ is the classifying topos for a geometric theory $\mathbb{T}$ over $\mathcal{S}$, then $B_U\mathcal{E}$ classifies the theory* $\mathbb{T}$-cobag. *In particular, the latter theory is geometric.*

(ii) *The functor B_U carries the structure of a* co-KZ-*monad on* $\mathfrak{BTop}/\mathcal{S}$.

Proof We shall only sketch the broad outline of the proof, leaving the reader to fill in the details.

(i) It is clear from the definition of $B_U\mathcal{E}$ as a partial product that geometric morphisms $\mathcal{F} \to B_U\mathcal{E}$ correspond to co-bags of $\mathbb{T}$-models in $\mathcal{F}$. To verify that geometric transformations between such morphisms correspond to morphisms of co-bags, we need to show that the contravariant functor

$\mathbb{B}(\mathcal{F}) \to \mathfrak{BTop}/\mathcal{F}$ which sends an internal Boolean algebra B in $\mathcal{F}$ to $\mathbf{Sh}_{\mathcal{F}}(B, C)$ is full and faithful (in the 2-categorical sense), i.e. that every morphism over $\mathcal{F}$ between entire $\mathcal{F}$-toposes is canonically isomorphic to one induced by a homomorphism between the corresponding Boolean algebras. This may be deduced by combining C1.4.5 with the fact that B is recoverable from $\mathbf{Sh}(B, C)$ as the lattice of complemented subterminal objects.

(ii) Given (i), this may be proved by means essentially similar to those by which we constructed the KZ-monad structure on B_L. The unit $\mathcal{E} \to B_U\mathcal{E}$ comes from regarding a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ as a co-bag (f, g) for which g is an identity morphism (it is clear that identity morphisms are entire, since if $2 = 1 \amalg 1$ denotes the initial Boolean algebra in $\mathcal{F}$ then $\mathbf{Sh}_{\mathcal{F}}(2, C) \simeq \mathcal{F}$). To obtain the multiplication, we similarly need to know that a composite of two entire morphisms is entire; this can be proved directly using the techniques of Section C2.5, but it also follows from the existence of the pure–entire factorization which we shall establish in C3.4.13. $\square$

We shall not attempt to give a general characterization of the algebras for the monad B_U, similar to 4.4.19, but we note one particular example of a free algebra:

Example 4.4.23 Let R be a commutative ring (in a topos $\mathcal{S}$, though for simplicity we shall think of $\mathcal{S}$ as **Set**). It is well known that the idempotents of R (that is, the elements e satisfying $e^2 = e$) form a Boolean algebra $E(R)$ in the ordering defined by $e \leq f$ iff $ef = e$. The *Pierce representation* of R expresses R as the ring of global sections of a sheaf of rings $\overline{R}$ on the Stone space of $E(R)$; explicitly, $\overline{R}$ corresponds to the sheaf for the coherent coverage on $E(R)$ whose sections over an idempotent e are the elements of the quotient ring $R/(1-e)$. We recall that a commutative ring is said to be *indecomposable* if it has exactly two idempotents 0 and 1, i.e. if it satisfies the coherent axioms $((0 = 1) \vdash \perp)$ and

$$((e^2 = e) \vdash_e ((e = 0) \vee (e = 1))) .$$

It is easy to verify that the Pierce sheaf $\overline{R}$, considered as a commutative ring in the topos $\mathbf{Sh}(E(R), C)$, satisfies these axioms (see [520, V 2.3]). Moreover, the universal property of the Pierce sheaf, described in [520, V 2.5], is precisely the statement that the specification of a commutative ring (in an arbitrary topos $\mathcal{F}$) is equivalent to the specification of a co-bag of idecomposable rings in $\mathcal{F}$. Thus we see that, if $\mathcal{E}$ is the classifying topos for the theory of indecomposable rings, $B_U\mathcal{E}$ is (equivalent to) the classifying topos for arbitrary commutative rings.

There are numerous other examples of algebraic structures which possess a 'Boolean representation theory' (cf. [290]) saying that arbitrary algebras are equivalent to co-bags of 'simple' algebras. In all such cases, we can argue as

in 4.4.23 to show that the classifying topos for the appropriate theory has the structure of an upper bagdomain.

Suggestions for further reading: Diers [290], Jibladze [499], Johnstone [535, 536, 537], Rosebrugh & Wood [1037, 1038, 1039, 1040], Street [1132], Vickers [1206].

B4.5 The symmetric monad

In this section we consider a close relative of the lower bagdomain construction of the previous section, introduced in [193] and extensively studied in [197] and [198], which throws light on the relationship between bounded $\mathcal{S}$-toposes and cocomplete $\mathcal{S}$-indexed categories. From 3.1.2, we know that every $\mathcal{S}$-topos is cocomplete and locally small as an $\mathcal{S}$-indexed category, and every inverse image functor over $\mathcal{S}$ is cocontinuous as an $\mathcal{S}$-indexed functor. We thus have a forgetful functor $(\mathfrak{Top}/\mathcal{S})^{\mathrm{op}} \to \mathfrak{Cocomp}_{\mathcal{S}}$, where $\mathfrak{Cocomp}_{\mathcal{S}}$ is the meta-2-category of cocomplete locally small $\mathcal{S}$-indexed categories, cocontinuous $\mathcal{S}$-indexed functors and indexed natural transformations between them.

The codomain of this functor is evidently much too large for us to hope that it might have any sort of adjoint: if we wish for the latter, we must impose further 'size restrictions' besides local smallness. However, it turns out that there is a very natural full sub-2-category of $\mathfrak{Cocomp}_{\mathcal{S}}$, which includes (the underlying indexed categories of) all bounded $\mathcal{S}$-toposes, and over which $(\mathfrak{BTop}/\mathcal{S})^{\mathrm{op}}$ may be regarded as the category of algebras for a KZ-monad. This subcategory is nothing other than the $\mathcal{S}$-indexed analogue of the 2-category of locally presentable categories in the sense of Gabriel and Ulmer [386], and we shall denote it by $\mathfrak{LocPres}_{\mathcal{S}}$. Locally presentable categories over **Set** are studied, in the more general context of accessible categories, in Section D2.3, and we shall see there that every Grothendieck topos is locally presentable as a category (D2.3.7). However, rather than translating the general definition of local presentability into $\mathcal{S}$-indexed terms, and proving the $\mathcal{S}$-indexed version of the result just cited, we shall here adopt a more *ad hoc* definition of what it means for a cocomplete $\mathcal{S}$-indexed category to be locally presentable. We shall make some remarks about the equivalence of this definition with the more usual one in 4.5.3 below; however, for the moment, readers who already know what locally presentable categories are would be well advised to forget that knowledge, and simply accept the fact that we have chosen $\mathfrak{LocPres}_{\mathcal{S}}$ as our (somewhat idiosyncratic) name for a class of 'well-behaved' cocomplete $\mathcal{S}$-indexed categories that we wish to study.

It will also be convenient to introduce a new name for the 2-category $(\mathfrak{BTop}/\mathcal{S})^{\mathrm{op}}$. Just as, in Chapter C1, we shall find it convenient to introduce different terminology and notation for the generalized spaces (locales) which we

study in that chapter and the algebraic entities (frames) which are extensionally the same thing, so it is sometimes convenient to distinguish terminologically between a bounded $\mathcal{S}$-topos, regarded as a 'generalized space over $\mathcal{S}$', and the $\mathcal{S}$-indexed category which is how we represent it extensionally, but which is really an algebraic gadget rather than a geometric one. Following a suggestion of S. J. Vickers [1210], we shall use the name *Giraud frame* (over $\mathcal{S}$) for an $\mathcal{S}$-indexed category $\mathbb{E}$ which satisfies the conditions of 3.3.11 (that is, derives from a bounded $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$), and write $\mathfrak{GFrm}_{\mathcal{S}}$ for the 2-category whose objects are Giraud frames over $\mathcal{S}$, whose 1-cells are cocontinuous cartesian $\mathcal{S}$-indexed functors (that is, inverse image functors), and whose 2-cells are $\mathcal{S}$-indexed natural transformations between such functors. (Thus $\mathfrak{GFrm}_{\mathcal{S}}$ is, up to equivalence, just another name for $(\mathfrak{BTop}/\mathcal{S})^{\mathrm{op}}$.)

Among the objects of $\mathfrak{LocPres}_{\mathcal{S}}$, we certainly wish to have all the diagram categories $[\mathbb{C}, \mathbb{S}]$, where $\mathbb{C}$ is an internal category in $\mathcal{S}$. We recall from 2.5.8 that $[\mathbb{C}, \mathbb{S}]$ is the free ($\mathcal{S}$-indexed) cocompletion of $\mathbb{C}^{\mathrm{op}}$, so that cocontinuous functors $[\mathbb{C}, \mathbb{S}] \to [\mathbb{D}, \mathbb{S}]$ correspond to arbitrary $\mathcal{S}$-indexed functors $\mathbb{C}^{\mathrm{op}} \to [\mathbb{D}, \mathbb{S}]$ – or equivalently, to profunctors $\mathbb{C} \nrightarrow \mathbb{D}$ in $\mathcal{S}$, as defined in 2.7.1. That is, we may regard the bicategory $\mathfrak{Prof}_{\mathcal{S}}$ of internal categories and profunctors as (equivalent to) a full sub-2-category of our desired category. We shall denote this subcategory by $\mathfrak{Dgm}_{\mathcal{S}}$, and call it the 2-category of diagram categories over $\mathcal{S}$.

But of course not every Giraud frame over $\mathcal{S}$ is of the form $[\mathbb{C}, \mathbb{S}]$: we need to add more objects to our 2-category. The surprising thing is that all we need to add are coinverters of 2-cells (cf. 1.1.4(*e*)): the key result that proves this is little more than a restatement of A4.3.11 (and also a close relative of 4.1.6).

Lemma 4.5.1 *Let $\mathcal{E}$ be a bounded $\mathcal{S}$-topos, and j a local operator on $\mathcal{E}$. Then there exists an internal category $\mathbb{D}$ in $\mathcal{S}$ and a 2-cell $\alpha\colon f \to g$ between morphisms $f, g\colon [\mathbb{D}, \mathcal{S}] \rightrightarrows \mathcal{E}$ in $\mathfrak{Cocomp}_{\mathcal{S}}$, such that the associated sheaf functor $i^*\colon \mathcal{E} \to \mathbf{sh}_j(\mathcal{E})$ has the universal property of a coinverter for α in $\mathfrak{Cocomp}_{\mathcal{S}}$.*

Proof We recall that an object B of $\mathcal{E}$ is a j-sheaf iff, for every j-dense monomorphism $A' \rightarrowtail A$ in $\mathcal{E}$, each morphism $A' \to B$ extends uniquely to a morphism $A \to B$. But it clearly suffices to consider only those dense monomorphisms $A' \rightarrowtail A$ for which A lies in some given separating family for $\mathcal{E}$; and the collection of such monomorphisms may be indexed by an object I of $\mathcal{S}$. We take $\mathbb{D}$ to be the discrete category indexed by I; then we have a pair of functors $\check{f}, \check{g}\colon \mathbb{D} \rightrightarrows \mathbb{E}$ and a natural transformation $\check{\alpha}$ between them which simply picks out this family of morphisms of $\mathcal{E}$. Since $\mathbb{E}$ is cocomplete as an $\mathcal{S}$-indexed category, $\check{\alpha}$ extends canonically to a natural transformation $\alpha\colon f \to g$ between cocontinuous functors $[\mathbb{D}, \mathcal{S}] \rightrightarrows \mathcal{E}$, by 2.5.8.

Since i^* inverts all j-dense monomorphisms, it is clear that $i^* \circ \alpha$ is an isomorphism. Now suppose $h\colon \mathbb{E} \to \mathbb{X}$ is any morphism of $\mathfrak{Cocomp}_{\mathcal{S}}$ such that $h \circ \alpha$ is an isomorphism. Since $\mathbb{E}$ is well-copowered and has a separating family, it follows from 2.4.6 that h has a right adjoint h_*; and the proof of A4.3.11 then shows that $h_*(X)$ is a j-sheaf for every object X of $\mathcal{X}$, i.e. h_* factors uniquely

through the right adjoint i_* of i^*. (In A4.3.11 we were dealing with adjunctions of which the left adjoint preserved finite limits, but that condition was not used in this part of the proof.) If we denote this factorization by k_*, it then follows easily from the fact that i_* is full and faithful that the composite $h \circ i_*$ is left adjoint to k_*, and hence that it is the unique factorization (up to canonical isomorphism) of h through i^*. □

Definition 4.5.2 We define the 2-category $\mathfrak{LocPres}_{\mathcal{S}}$ of *locally presentable $\mathcal{S}$-indexed categories* to have as objects all those cocomplete locally small $\mathcal{S}$-indexed categories $\mathbb{X}$ which can be presented as the codomain of a coinverter

$$\begin{tikzcd} {[\mathbb{D},\mathbb{S}]} \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & {[\mathbb{C},\mathbb{S}]} \arrow[r] & \mathbb{X} \end{tikzcd}$$

in $\mathfrak{Cocomp}_{\mathcal{S}}$, together with all $\mathcal{S}$-cocontinuous functors between them and all $\mathcal{S}$-indexed natural transformations between the latter.

Since every bounded $\mathcal{S}$-topos can be presented as a subtopos of one of the form $[\mathbb{C}, \mathcal{S}]$, it follows immediately from 4.5.1 that the forgetful functor $\mathfrak{GFrm}_{\mathcal{S}} \to \mathfrak{Cocomp}_{\mathcal{S}}$ factors through $\mathfrak{LocPres}_{\mathcal{S}}$.

We recall that in 4.1.7 we showed that arbitrary inverters exist in $\mathfrak{BTop}/\mathcal{S}$; equivalently, $\mathfrak{GFrm}_{\mathcal{S}}$ has coinverters. And a comparison of that proof with 4.5.1 above will show that the forgetful functor $\mathfrak{GFrm}_{\mathcal{S}} \to \mathfrak{LocPres}_{\mathcal{S}}$ preserves coinverters. Thus, if the functors f and g appearing in the displayed diagram in 4.5.2 are actually inverse image functors (i.e. if they are cartesian), then $\mathcal{X}$ is actually a bounded $\mathcal{S}$-topos.

Remark 4.5.3 Classically, a locally presentable category (over **Set**) is one which is representable as the category of κ-continuous **Set**-valued functors on a κ-complete small category $\mathcal{C}$, where κ is a cardinal and 'κ-complete' (resp. 'κ-continuous') means 'having (resp. preserving) limits of diagrams of cardinality less than κ'. (cf. D2.3.4). By a similar argument to that in the proof of 4.5.1, we may represent it as an inverter (in the 2-category $\mathfrak{Comp}_{\mathbf{Set}}$ of complete locally small categories and continuous functors) of a 2-cell between functors $[\mathcal{C}, \mathbf{Set}] \rightrightarrows [\mathcal{D}, \mathbf{Set}]$ (where $\mathcal{D}$ may in fact be taken to be discrete); but since any continuous functor between locally presentable categories has a left adjoint, by the Special Adjoint Functor Theorem, we may take left adjoints of the functors concerned to obtain a representation of our locally presentable category as a coinverter in $\mathfrak{Cocomp}_{\mathbf{Set}}$. On the other hand, a theorem of Makkai and Paré [787] asserts that the 2-category of locally presentable categories is closed under weighted limits (in particular, under inverters) in $\mathfrak{Comp}_{\mathbf{Set}}$; so we may identify the locally presentable categories of 4.5.2 (in the case $\mathcal{S} = \mathbf{Set}$) with those defined in the traditional way.

We now embark on our study of the relationship between $\mathfrak{GFrm}_S$ and $\mathfrak{LocPres}_S$. From this point on, we shall need to assume that our base topos $\mathcal{S}$ has a natural number object.

Proposition 4.5.4 *Let $\mathcal{S}$ be a topos with a natural number object. Then the forgetful functor $\mathfrak{GFrm}_S \to \mathfrak{LocPres}_S$ has a left adjoint Σ.*

Given a locally presentable $\mathcal{S}$-indexed category $\mathbb{C}$, we call $\Sigma\mathbb{C}$ the *symmetric Giraud frame* (or symmetric topos) generated by $\mathcal{E}$. The reason for the name is an analogy with classical commutative algebra, obtained by thinking of finite limits as 'multiplication' and $\mathcal{S}$-indexed colimits as 'addition' and 'multiplication by scalars in $\mathcal{S}$': in this analogy, cartesian categories are 'commutative monoids', cocomplete $\mathcal{S}$-indexed categories are 'modules', and Giraud frames are 'algebras'. Thus the free functor Σ is the analogue of the symmetric algebra construction. (The analogy becomes rather more explicit when we compare (ordinary) frames with complete join-semilattices, as we shall see in Section C1.1.)

Proof First let $\mathbb{C}$ be an internal category in $\mathcal{S}$. We know from 2.5.8 that $[\mathbb{C}, \mathbb{S}]$ is the free $\mathcal{S}$-indexed cocompletion of $\mathbb{C}^{\mathrm{op}}$: that is, for any cocomplete $\mathcal{S}$-indexed category $\mathbb{E}$, arbitrary $\mathcal{S}$-indexed functors $\mathbb{C}^{\mathrm{op}} \to \mathbb{E}$ correspond to cocontinuous indexed functors $[\mathbb{C}, \mathbb{S}] \to \mathbb{E}$. But we also know from 3.2.6 that we may construct the free cocartesian category $\mathrm{CoCart}(\mathbb{C})$ on $\mathbb{C}$ as an internal category in $\mathcal{S}$; so if $\mathbb{E}$ is a Giraud frame then arbitrary functors $\mathbb{C}^{\mathrm{op}} \to \mathbb{E}$ correspond to cartesian functors $\mathrm{CoCart}(\mathbb{C})^{\mathrm{op}} \to \mathbb{E}$, which in turn correspond by 3.2.7 to inverse image functors $[\mathrm{CoCart}(\mathbb{C}), \mathbb{S}] \to \mathbb{E}$. Thus we have shown that $\Sigma[\mathbb{C}, \mathbb{S}]$ may be taken to be the Giraud frame $[\mathrm{CoCart}(\mathbb{C}), \mathbb{S}]$. Moreover, it is clear that this definition of Σ extends to a 2-functor on $\mathfrak{Dgm}_S$: given a cocontinuous functor $f\colon [\mathbb{C}, \mathbb{S}] \to [\mathbb{D}, \mathbb{S}]$, we may regard it as a functor $\mathbb{C}^{\mathrm{op}} \to [\mathbb{D}, \mathbb{S}]$, compose it with the unit map $[\mathbb{D}, \mathbb{S}] \to \Sigma[\mathbb{D}, \mathbb{S}]$ (which is simply left Kan extension along the inclusion $\mathbb{D} \to \mathrm{CoCart}(\mathbb{D})$), and then extend it to a cartesian functor $\mathrm{CoCart}(\mathbb{C})^{\mathrm{op}} \to \Sigma[\mathbb{D}, \mathbb{S}]$, which in turn defines a morphism of Giraud frames $\Sigma[\mathbb{C}, \mathbb{S}] \to \Sigma[\mathbb{D}, \mathbb{S}]$. The argument for 2-cells is similar.

Defining Σ in the general case is now easy, since, as a left adjoint, it must preserve coinverters. Given a coinverter presentation

$$\begin{tikzcd} {[\mathbb{D}, \mathbb{S}]} \arrow[r, shift left=2, "f"] \arrow[r, shift right=2, "g"'] \arrow[r, phantom, "\Downarrow\alpha"] & {[\mathbb{C}, \mathbb{S}]} \arrow[r, "h"] & \mathbb{X} \end{tikzcd}$$

of an object $\mathbb{X}$ of $\mathfrak{LocPres}_S$, we form the corresponding 2-cell

$$\begin{tikzcd} {[\mathrm{CoCart}(\mathbb{D}), \mathbb{S}]} \arrow[r, shift left=2, "f^*"] \arrow[r, shift right=2, "g^*"'] \arrow[r, phantom, "\Downarrow\alpha^*"] & {[\mathrm{CoCart}(\mathbb{C}), \mathbb{S}]} \end{tikzcd}$$

in $\mathfrak{GFrm}_S$, and take its coinverter in this 2-category (equivalently, the inverter in $\mathfrak{BTop}/S$, cf. 4.1.7) to define $\Sigma\mathbb{X}$. □

Now consider the diagram

$$\begin{tikzcd}
{[\mathbb{D},\mathbb{S}]} \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] \arrow[r, phantom, "\downarrow\alpha"] \arrow[d, "\eta_{[\mathbb{D},\mathbb{S}]}"] & {[\mathbb{C},\mathbb{S}]} \arrow[r, "h"] \arrow[d, "\eta_{[\mathbb{C},\mathbb{S}]}"] & \mathbb{X} \arrow[d, "\eta_{\mathbb{X}}"] \\
{[\mathrm{CoCart}(\mathbb{D}),\mathbb{S}]} \arrow[r, shift left, "f^*"] \arrow[r, shift right, "g_*"'] \arrow[r, phantom, "\downarrow\alpha^*"] & {[\mathrm{CoCart}(\mathbb{C}),\mathbb{S}]} \arrow[r, "h^*"] & \Sigma\mathbb{X}
\end{tikzcd}$$

associated with a coinverter presentation of an object $\mathbb{X}$ of $\mathfrak{LocPres}_S$, where the vertical arrows are the unit maps. As we remarked in the proof of Proposition 4.5.4, $\eta_{[\mathbb{C},\mathbb{S}]}$ is the left Kan extension functor along the full embedding $i\colon \mathbb{C} \to \mathrm{CoCart}(\mathbb{C})$; so its right adjoint $\theta_{\mathbb{C}}$ (which exists as an indexed functor, simply because η is cocontinuous) is itself a morphism of $\mathfrak{LocPres}_S$; also, the unit of the adjunction $(\eta \dashv \theta)$ is an isomorphism. More generally, we have

Theorem 4.5.5 *Let S be a topos with a natural number object, and $\mathbb{X}$ an object of $\mathfrak{LocPres}_S$. Then $\mathbb{X}$ is a Giraud frame iff $\eta_{\mathbb{X}}\colon \mathbb{X} \to \Sigma\mathbb{X}$ has a right adjoint in $\mathfrak{LocPres}_S$, such that the unit of the adjunction is an isomorphism.*

Proof First suppose $\mathbb{X}$ is a Giraud frame. Then, as we saw in 4.5.1, we may choose its coinverter presentation so that $h\colon [\mathbb{C},\mathbb{S}] \to \mathbb{X}$ is actually a morphism of $\mathfrak{GFrm}_S$. But $\theta_{\mathbb{C}}\colon [\mathrm{CoCart}(\mathbb{C}),\mathbb{S}] \to [\mathbb{C},\mathbb{S}]$ is also a morphism of $\mathfrak{GFrm}_S$, since it has adjoints on both sides; and $h\theta_{\mathbb{C}}$ coinverts the 2-cell α^*, since

$$h\theta_{\mathbb{C}} \circ \alpha^* \circ \eta_{[\mathbb{D},\mathbb{S}]} = h\theta_{\mathbb{C}}\eta_{[\mathbb{C},\mathbb{S}]} \circ \alpha = h \circ \alpha$$

which is an isomorphism, so we get a factorization $\theta_{\mathbb{X}}\colon \Sigma\mathbb{X} \to \mathbb{X}$ of $h\theta_{\mathbb{C}}$ through h^*. Similar arguments show that the composite $\theta_{\mathbb{X}}\eta_{\mathbb{X}}$ is isomorphic to the identity, and that the counit 2-cell from $\eta_{[\mathbb{C},\mathbb{S}]}\theta_{\mathbb{C}}$ to the identity factors through h^* to yield a 2-cell which is the counit of an adjunction $(\eta_{\mathbb{X}} \dashv \theta_{\mathbb{X}})$, the unit being the isomorphism already mentioned.

Conversely, suppose $\eta_{\mathbb{X}}$ has a right adjoint θ in $\mathfrak{LocPres}_S$ such that the unit of the adjunction is an isomorphism. Since $\Sigma_{\mathbb{X}}$ is a Giraud frame, we can use the Special Adjoint Functor Theorem (2.4.6) to obtain a further right adjoint θ_* for θ (though the latter will not, of course, be a morphism of $\mathfrak{LocPres}_S$ in general); and from the fact that the unit of $(\eta_{\mathbb{X}} \dashv \theta)$ is an isomorphism, it follows that the counit of $(\theta \dashv \theta_*)$ is an isomorphism. Also, θ is cartesian since it has a left adjoint; and so we have identified $\mathcal{X}$ (the underlying ordinary category of $\mathbb{X}$) with a reflective subcategory of $\Sigma\mathcal{X}$ for which the reflector is cartesian. Hence by A4.3.9 it is a topos; so $\mathbb{X}$ is a Giraud frame. □

the smallest local operator on $M\mathcal{E}$ whose associated sheaf functor maps $\delta_!(1_{\mathcal{E}}) \to 1_{M\mathcal{E}}$ to an isomorphism (cf. A4.5.14(c)). Since the unit map $1_{\mathcal{E}} \to \delta^*\delta_!(1_{\mathcal{E}})$ is an isomorphism, it is clear that δ itself factors through $P\mathcal{E} \to M\mathcal{E}$. More generally, a geometric morphism $f\colon \mathcal{F} \to M\mathcal{E}$ over $\mathcal{S}$ factors through $P\mathcal{E}$ iff the cocontinuous $\mathcal{S}$-indexed functor $\mathbb{E} \to \mathbb{F}$ to which it corresponds (which is, of course, just the canonical $\mathcal{S}$-indexing of the composite $f^*\delta_!$) preserves 1, i.e. it is an $\mathcal{F}$-valued probability measure on $\mathcal{E}$.

If $\mathcal{E} = [\mathbb{C}, \mathcal{S}]$ for an internal category $\mathbb{C}$, then $P\mathcal{E}$ may be identified with $[\mathrm{Conn}_f(\mathbb{C}), \mathcal{S}]$, where $\mathrm{Conn}_f(\mathbb{C})$ denotes the free completion of $\mathbb{C}$ with respect to connected finite colimits. To see this, we do not need to give an explicit construction of $\mathrm{Conn}_f(\mathbb{C})$; but we note that, as usual, it may be identified with the full subcategory of $[\mathbb{C}^{\mathrm{op}}, \mathbb{S}]$ whose objects are connected finite colimits of representables, from which it is easy to see that the full embedding $\mathbb{C} \to \mathrm{Conn}_f(\mathbb{C})$ is an initial functor. Now a cocontinuous $\mathcal{S}$-indexed functor $[\mathbb{C}, \mathbb{S}] \to \mathbb{F}$ preserves 1 (that is, it is a probability measure on $[\mathbb{C}, \mathcal{S}]$) iff the diagram of shape $\mathbb{C}^{\mathrm{op}}$ in $\mathcal{F}$ which is its restriction to representables (cf. 2.5.8) has colimit 1; so by 2.5.12 we deduce that the same is true of its unique connected-finite-limit-preserving extension to a functor $\mathrm{Conn}_f(\mathbb{C})^{\mathrm{op}} \to \mathbb{F}$. But the preservation of finite connected limits by the latter easily implies, as in 3.2.5, that the total category of the corresponding discrete opfibration is weakly filtered; hence by 2.6.5 it is filtered. So it is a $\mathrm{Conn}_f(\mathbb{C})$-torsor, and hence it induces a geometric morphism $\mathcal{F} \to [\mathrm{Conn}_f(\mathbb{C}), \mathcal{S}]$. The argument in the opposite direction is similar. It follows that, for any $\mathcal{E}$, we can construct a presentation of $P\mathcal{E}$ as an inverter in $\mathfrak{BTop}/\mathcal{S}$ from such a presentation of $\mathcal{E}$, in the same way as we did for $M\mathcal{E}$ but replacing the functor $\mathrm{CoCart}(-)$ by $\mathrm{Conn}_f(-)$.

We next claim that, for any $\mathcal{E}$, we have $B_L(P\mathcal{E}) \simeq M\mathcal{E}$. Again, this could be proved by first considering toposes of the form $[\mathbb{C}, \mathcal{S}]$, and noting that we have $\mathrm{Fam}_f(\mathrm{Conn}_f(\mathbb{C})) \simeq \mathrm{CoCart}(\mathbb{C})$, since any finite colimit can be uniquely decomposed as a finite coproduct of connected colimits. But we can derive the same result more simply, by noting that any cocontinuous $\mathcal{S}$-indexed functor $f\colon \mathbb{E} \to \mathbb{F}$ can be factored as

$$\mathbb{E} \longrightarrow \mathbb{F}/f(1) \longrightarrow \mathbb{F}$$

(just as in A1.2.9, but without the hypothesis that f preserves pullbacks) where the first factor is cocontinuous and preserves the terminal object. Conversely, given an object B of $\mathcal{F}$ and a cocontinuous 1-preserving functor $\mathbb{E} \to \mathbb{F}/B$, we may compose with the forgetful functor to obtain a cocontinuous functor $\mathbb{E} \to \mathbb{F}$. So $\mathcal{F}$-valued measures on $\mathcal{E}$ are essentially the same thing as bags of $\mathcal{F}$-valued probability measures on $\mathcal{E}$; hence the toposes which represent the two notions must be equivalent.

On the other hand, we do not in general have $P(B_L\mathcal{E}) \simeq M\mathcal{E}$, since we do not have $\mathrm{Conn}_f(\mathrm{Fam}_f(\mathbb{C})) \simeq \mathrm{CoCart}(\mathbb{C})$; although any finite colimit can be written as a connected colimit of finite coproducts (for example, as a coequalizer

Corollary 4.5.6 *The monad induced by the adjunction between the forgetful functor $\mathfrak{GFrm}_S \to \mathfrak{LocPres}_S$ and its left adjoint Σ is a co-KZ-monad, and the adjunction is monadic.*

Proof The construction of the right adjoint to $\eta_{\mathbb{X}}$, for a Giraud frame $\mathbb{X}$, in fact shows that it is the component of the counit of the adjunction at the object $\mathbb{X}$; in particular, if $\mathbb{X}$ is itself of the form $\Sigma\mathbb{Y}$, then it is the component of the multiplication of the monad at $\mathbb{Y}$. Hence the multiplication is (naturally) right adjoint to the unit of the monad. Also, we have shown in 4.5.5 that the objects of $\mathfrak{LocPres}_S$ which admit algebra structures for the monad are exactly the Giraud frames (cf. 1.1.13). So, to complete the proof, we have only to show that if $\mathbb{E}$ and $\mathbb{F}$ are Giraud frames then a cocontinuous functor $f\colon \mathbb{E} \to \mathbb{F}$ is a morphism of Σ-algebras if it is cartesian. But if f is a morphism of Σ-algebras, then the composite $f\theta_{\mathbb{E}} \cong \theta_{\mathbb{F}} f^*$ is cartesian, whence f is cartesian since $\theta_{\mathbb{E}}$ is a coreflector. Conversely, if f is cartesian, then it is a morphism of $\mathfrak{GFrm}_S$, and so induces a morphism of Σ-algebras (namely, itself!) via the comparison functor. □

Next, we note that the construction of $\Sigma\mathbb{X}$ may be factored into two stages. Let us, for want of a better name, write $\mathfrak{K}$ for the full sub-2-category of $\mathfrak{LocPres}_S$ whose objects are the Giraud frames; thus we have inclusions

$$\mathfrak{GFrm}_S \rightarrowtail \mathfrak{K} \rightarrowtail \mathfrak{LocPres}_S$$

of which the first is bijective on objects and the second is full and faithful.

Proposition 4.5.7 *The inclusion $\mathfrak{K} \to \mathfrak{LocPres}_S$ has a left adjoint T. Moreover, for every object $\mathbb{X}$ of $\mathfrak{LocPres}_S$, we have $\Sigma\mathbb{X} \simeq \Sigma T\mathbb{X}$.*

Proof Given $\mathbb{X}$, we may present it as usual as a coinverter

$$\begin{tikzcd}
{[\mathbb{D},\mathbb{S}]} \arrow[r, shift left=2, "f"] \arrow[r, shift right=2, "g"'] \arrow[r, phantom, "\downarrow\alpha"] & {[\mathbb{C},\mathbb{S}]} \arrow[r, "h"] & \mathbb{X}
\end{tikzcd}$$

and consider the pushout

$$\begin{tikzcd}
{[\mathrm{CoCart}(\mathbb{C}),\mathbb{S}]} \arrow[r, "h^*"] \arrow[d, "\theta_{\mathbb{C}}"] & \Sigma_{\mathbb{X}} \arrow[d, "t"] \\
{[\mathbb{C},\mathbb{S}]} \arrow[r, "h^+"] & T\mathbb{X}
\end{tikzcd}$$

in $\mathfrak{GFrm}_S$, i.e. the pullback in $\mathfrak{BTop}/S$. From the construction of pullbacks of inclusions in 3.3.6, it is easily verified (by an argument like that in the proof of

4.5.1) that the above square is also a pushout in $\mathfrak{LocPres}_{\mathcal{S}}$, and that h^+ is the coinverter in $\mathfrak{GFrm}_{\mathcal{S}}$ (or in $\mathfrak{LocPres}_{\mathcal{S}}$) of the 2-cell

$$\begin{tikzcd}[\mathrm{CoCart}(\mathbb{D}), \mathbb{S}] \arrow[r, bend left, "f^+"] \arrow[r, bend right, "g^+"'] \arrow[r, phantom, "\downarrow \alpha^+"] & [\mathbb{C}, \mathbb{S}]\end{tikzcd}$$

induced by α in the obvious way. Hence in particular there is a unique (up to isomorphism) cocontinuous functor $u \colon T\mathbb{X} \to \Sigma\mathbb{X}$ satisfying $uh^+ \cong h^*\eta_{[\mathbb{C},\mathbb{S}]}$ (u is in fact the left adjoint of the morphism t in the pushout diagram, though we don't need this).

Since $\alpha^+ \circ \eta_{[\mathbb{D},\mathbb{S}]} = \alpha$, it is clear that we get a canonical morphism $\mathbb{X} \to T\mathbb{X}$ in $\mathfrak{LocPres}_{\mathcal{S}}$. Also, if $k \colon \mathbb{X} \to \mathbb{E}$ is any cocontinuous functor from $\mathbb{X}$ to a Giraud frame, then it extends uniquely to a Giraud frame morphism $k^* \colon \Sigma\mathbb{X} \to \mathbb{E}$, and now we have

$$k^*uth^* \cong k^*uh^+\theta_{\mathbb{C}} \cong k^*h^*\eta_{[\mathbb{C},\mathbb{S}]}\theta_{\mathbb{C}} \cong k^*\eta_{\mathbb{X}}h\theta_{\mathbb{C}} \cong kh\theta_{\mathbb{C}}$$

so that the pair (k^*ut, kh) induces a unique cocontinuous functor $k^+ \colon T\mathbb{X} \to \mathbb{E}$ (which must in fact be the composite k^*u). Thus we have shown that $T\mathbb{X}$ is the required left adjoint. Also, since for any Giraud frame $\mathbb{E}$ the categories of cocontinuous functors $\mathbb{X} \to \mathbb{E}$ and $T\mathbb{X} \to \mathbb{E}$ are equivalent, we see that $\Sigma\mathbb{X}$ and $\Sigma T\mathbb{X}$ have the same universal property in $\mathfrak{GFrm}_{\mathcal{S}}$, and so must be equivalent.

□

Thus, if we wish to characterize the 'symmetric Giraud frames' (that is, those which occur as $\Sigma\mathbb{X}$ for some $\mathbb{X}$), it suffices to consider those of the form $\Sigma\mathbb{E}$ where $\mathbb{E}$ is already a Giraud frame. In order to do this, we shall investigate the comonad on $\mathfrak{GFrm}_{\mathcal{S}}$ induced by Σ and its right adjoint; but at this point it will be convenient to reverse our perspective to the more familiar 'geometric' one, and regard the structure in question as a monad on $\mathfrak{BTop}/\mathcal{S}$. Rotating the symbol Σ through 90° rather than 180° as might have been expected, we shall denóte the bounded $\mathcal{S}$-topos corresponding to the Giraud frame $\Sigma\mathbb{E}$ by $M\mathcal{E}$. (The reason for this notation is that, as suggested by F. W. Lawvere, it is helpful to regard the points of $M\mathcal{E}$ – equivalently, the cocontinuous $\mathcal{S}$-indexed functors $\mathbb{E} \to \mathbb{S}$ – as 'measures' or 'distributions' on $\mathcal{E}$, those which derive from points of $\mathcal{E}$ – equivalently, which correspond to cartesian cocontinuous functors – being the 'Dirac measures' concentrated at a single point. We shall make this idea more precise in 4.5.9 below.)

The unit of the monad on $\mathfrak{BTop}/\mathcal{S}$ is the geometric morphism $\delta_{\mathcal{E}} \colon \mathcal{E} \to M\mathcal{E}$ whose inverse image is the Σ-algebra structure $\theta \colon \Sigma\mathbb{E} \to \mathbb{E}$; note that this morphism is an $\mathcal{S}$-essential inclusion in the sense of 3.1.1, since θ has an $\mathcal{S}$-indexed left adjoint $\eta_{\mathbb{E}}$ which is full and faithful. The multiplication $MM\mathcal{E} \to M\mathcal{E}$ is similarly the morphism whose inverse image is $\Sigma\eta_{\mathbb{E}} \colon \Sigma\mathbb{E} \to \Sigma\Sigma\mathbb{E}$; by 1.1.12, thi

Recall that we saw in 4.4.19 that B_L-algebras have a similar property with arbitrary $\mathcal{S}$-indexed colimits replaced by coproducts; and indeed $B_L\mathcal{E}$ is the free coproduct-completion of $\mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$. It is not true in general that $M\mathcal{E}$ is the free cocompletion of $\mathcal{E}$ in a similar 'external' sense. For a counterexample, take $\mathcal{E}$ to be a bounded **Set**-topos such that the geometric morphism $\gamma\colon \mathcal{E} \to \mathbf{Set}$ is essential, but $\mathcal{E}$ has no points (we shall give an example of such a topos in D3.4.14). Then $\mathfrak{Top}\,(\mathbf{Set}, \mathcal{E})$ is empty, but $\mathfrak{Top}\,(\mathbf{Set}, M\mathcal{E})$ contains (the measures corresponding to) all set-indexed copowers of $\gamma_!$, so it is not the free cocompletion of $\mathfrak{Top}\,(\mathbf{Set}, \mathcal{E})$. However, we shall see in 4.5.13 below that $M\mathcal{E}$ is the free $\mathcal{A}$-cocompletion of $\mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$ in the 'intrinsic' sense introduced in 1.1.16, for a suitable class $\mathcal{A}$ of geometric morphisms.

We certainly have a comparison map $B_L\mathcal{E} \to M\mathcal{E}$ for any $\mathcal{E}$: for we have a geometric morphism $B_L\mathcal{E}/I \to \mathcal{E}$ over $\mathcal{S}$, where I is the object of $B_L\mathcal{E}$ indexing the generic bag of points of $\mathcal{E}$, and if we compose the inverse image of this with the forgetful functor $B_L\mathcal{E}/I \to B_L\mathcal{E}$ we obtain a cocontinuous $\mathcal{S}$-indexed functor $\mathbb{E} \to B_L\mathbb{E}$, corresponding to a Giraud frame homomorphism $\Sigma\mathbb{E} \to B_L\mathbb{E}$.

In the case when $\mathcal{E}$ is of the form $[\mathbb{C}, \mathcal{S}]$, the comparison may be described more simply. In this case, we know that $B_L\mathcal{E}$ is the topos $[\mathrm{Fam}_f(\mathbb{C}), \mathcal{S}]$, where $\mathrm{Fam}_f(\mathbb{C})$ as usual denotes the free finite-coproduct-completion of $\mathbb{C}$ (cf. 4.4.17), and $M\mathcal{E}$ is $[\mathrm{CoCart}(\mathbb{C}), \mathcal{S}]$. Since the full embedding $\mathbb{C} \to \mathrm{CoCart}(\mathbb{C})$ obviously factors through $\mathrm{Fam}_f(\mathbb{C})$, we obtain a corresponding factorization

$$\mathcal{E} \longrightarrow B_L\mathcal{E} \longrightarrow M\mathcal{E}$$

of the unit $\delta_{\mathcal{E}}\colon \mathcal{E} \to M\mathcal{E}$, where both factors are essential inclusions. From the universal property of the lower bagdomain, it is easy to verify that the functor B_L preserves inverters in $\mathfrak{BTop}/\mathcal{S}$; hence, by taking a presentation of an arbitrary bounded $\mathcal{S}$-topos $\mathcal{E}$ as an inverter of a 2-cell between diagram toposes (cf. 4.5.1), we may deduce the existence of the corresponding factorization of $\delta_{\mathcal{E}}$, for an arbitrary $\mathcal{E}$. In particular, we deduce that the natural map $B_L\mathcal{E} \to M\mathcal{E}$ is an inclusion for any $\mathcal{E}$. In fact it is an essential inclusion: the left adjoint of its inverse image may be obtained by composing the left adjoint of the inverse image of the inclusion $B_L\mathcal{E} \to MB_L\mathcal{E}$ with the inverse image of $M(\mathcal{E} \to B_L\mathcal{E})$.

But there is another way in which we may factor $\delta_{\mathcal{E}}$. We noted in 4.4.15 that $B_L\mathcal{E}$ is a representing object for pre-geometric morphisms with codomain $\mathcal{E}$; that is, those morphisms which we get by forgetting the requirement that an inverse image functor must preserve the terminal object, but remembering that it must preserve pullbacks. Similarly, $M\mathcal{E}$ represents those $\mathcal{S}$-indexed adjunctions with codomain $\mathcal{E}$ which we get by forgetting altogether the requirement that an inverse image functor should preserve finite limits. If we forget that it should preserve pullbacks, but remember that it should preserve the terminal object, we arrive at the notion of a *probability measure* on $\mathcal{E}$, that is one for which 'the total measure of $\mathcal{E}$ is unity'. This too corresponds to a KZ-monad on $\mathfrak{BTop}/\mathcal{S}$; we may obtain its functor part as follows. Let $P\mathcal{E}$ be the subtopos of $M\mathcal{E}$ corresponding to

is right adjoint to $\theta_{\Sigma\mathbb{E}}$ in $\mathfrak{GFrm}_{\mathcal{S}}$, but because of the reversal of 1-cells it becomes *left* adjoint to the unit in $\mathfrak{BTop}/\mathcal{S}$. Thus we have verified

Lemma 4.5.8 *The monad on $\mathfrak{BTop}/\mathcal{S}$ whose functor part is M is a KZ-monad.* □

Remark 4.5.9 As we noted in the proof of 3.2.9, the free cocartesian category CoCart(**1**) on the terminal object of **Cat**($\mathcal{S}$) is (equivalent to) $\mathbb{S}_f$; so we see that $M\mathcal{S} \simeq [\mathbb{S}_f, \mathcal{S}]$ is the object classifier $\mathcal{S}[\mathbb{O}]$ over $\mathcal{S}$. In particular, $\mathcal{S}[\mathbb{O}]$ carries an M-algebra structure. In terms of the identification of points of $M\mathcal{E}$ with 'measures on $\mathcal{E}$', the algebra structure map $\int : M(\mathcal{S}[\mathbb{O}]) \to \mathcal{S}[\mathbb{O}]$ may be thought of as 'integration', in the following sense. We may think of an object A of $\mathcal{E}$ as a 'continuous $\mathcal{S}$-valued function' on the 'generalized space' $\mathcal{E}$ (this is an appealing way of viewing sheaves on traditional spaces), whose value at a point $s: \mathcal{S} \to \mathcal{E}$ is the object $s^*(A)$. Given a measure on $\mathcal{E}$, that is a point $m: \mathcal{S} \to M\mathcal{E}$, the 'integral' of A with respect to the measure is the object of $\mathcal{S}$ classified by the composite

$$\mathcal{S} \xrightarrow{\quad m \quad} M\mathcal{E} \xrightarrow{\quad M(\tilde{A}) \quad} M(\mathcal{S}[\mathbb{O}]) \xrightarrow{\quad \int \quad} \mathcal{S}[\mathbb{O}],$$

where $\tilde{A}$ is the classifying map of A. It is straightforward to verify that this is indeed the value at A of the cocontinuous functor $\mathbb{E} \to \mathbb{S}$ corresponding to m; in particular, if $m = \delta_{\mathcal{E}} s$ is the Dirac measure corresponding to a point s of $\mathcal{E}$, it is just $s^*(A)$.

Example 4.5.10 Among examples of M-algebras, we have all toposes of the form $[\mathbb{C}, \mathcal{S}]$ where $\mathbb{C}$ is cocartesian; for if $\mathbb{C}$ is cocartesian then the embedding $\mathbb{C} \to \text{CoCart}(\mathbb{C})$ has a left adjoint, given by forming colimits in $\mathbb{C}$, and hence $[\mathbb{C}, \mathcal{S}]$ is an adjoint retract of $[\text{CoCart}(\mathbb{C}), \mathcal{S}] \simeq M[\mathbb{C}, \mathcal{S}]$. Hence also any retract in $\mathfrak{BTop}/\mathcal{S}$ of a topos of the form $[\mathbb{C}, \mathcal{S}]$ with $\mathbb{C}$ cocartesian admits an M-algebra structure. (We shall study such toposes in greater detail in Section C4.3.)

Proposition 4.5.11 *If a bounded $\mathcal{S}$-topos $\mathcal{E}$ has the structure of an M-algebra, then for any bounded $\mathcal{S}$-topos $\mathcal{F}$ the $\mathcal{S}$-indexed category $\mathfrak{BTop}/\mathcal{S}\,(\mathbb{F}, \mathbb{E})$ (that is, the category whose I-indexed families are morphisms $\mathcal{F}/I \to \mathcal{E}/I$ over $\mathcal{S}/I$, for each object I of $\mathcal{S}$) is $\mathcal{S}$-cocomplete. Moreover, for any geometric morphism $f: \mathcal{F}' \to \mathcal{F}$ the functor $\mathfrak{BTop}/\mathcal{S}\,(\mathbb{F}, \mathbb{E}) \to \mathfrak{BTop}/\mathcal{S}\,(\mathbb{F}', \mathbb{E})$ induced by composition with f is cocontinuous, and similarly for the functor induced by composition with any M-algebra homomorphism $g: \mathcal{E} \to \mathcal{E}'$.*

Proof If $\mathcal{E} \simeq M\mathcal{G}$ is a free M-algebra, then $\mathfrak{BTop}/\mathcal{S}\,(\mathbb{F}, \mathbb{E})$ is equivalent to the category $\mathfrak{LocPres}_{\mathcal{S}}\,(\mathbb{G}, \mathbb{F})$, which is clearly cocomplete since it it closed under (pointwise) colimits in the category of all indexed functors $\mathbb{G} \to \mathbb{F}$. In general, any M-algebra $\mathcal{E}$ is a retract of a free M-algebra by 1.1.13, so the cocompleteness result follows from 1.1.10. The cocontinuity assertions are similarly straightforward. □

Recall that we saw in 4.4.19 that B_L-algebras have a similar property with arbitrary $\mathcal{S}$-indexed colimits replaced by coproducts; and indeed $B_L\mathcal{E}$ is the free coproduct-completion of $\mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$. It is not true in general that $M\mathcal{E}$ is the free cocompletion of $\mathcal{E}$ in a similar 'external' sense. For a counterexample, take $\mathcal{E}$ to be a bounded **Set**-topos such that the geometric morphism $\gamma\colon \mathcal{E} \to \mathbf{Set}$ is essential, but $\mathcal{E}$ has no points (we shall give an example of such a topos in D3.4.14). Then $\mathfrak{Top}(\mathbf{Set}, \mathcal{E})$ is empty, but $\mathfrak{Top}(\mathbf{Set}, M\mathcal{E})$ contains (the measures corresponding to) all set-indexed copowers of $\gamma_!$, so it is not the free cocompletion of $\mathfrak{Top}(\mathbf{Set}, \mathcal{E})$. However, we shall see in 4.5.13 below that $M\mathcal{E}$ is the free $\mathcal{A}$-cocompletion of $\mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$ in the 'intrinsic' sense introduced in 1.1.16, for a suitable class $\mathcal{A}$ of geometric morphisms.

We certainly have a comparison map $B_L\mathcal{E} \to M\mathcal{E}$ for any $\mathcal{E}$: for we have a geometric morphism $B_L\mathcal{E}/I \to \mathcal{E}$ over $\mathcal{S}$, where I is the object of $B_L\mathcal{E}$ indexing the generic bag of points of $\mathcal{E}$, and if we compose the inverse image of this with the forgetful functor $B_L\mathcal{E}/I \to B_L\mathcal{E}$ we obtain a cocontinuous $\mathcal{S}$-indexed functor $\mathbb{E} \to B_L\mathbb{E}$, corresponding to a Giraud frame homomorphism $\Sigma\mathbb{E} \to B_L\mathbb{E}$.

In the case when $\mathcal{E}$ is of the form $[\mathbb{C}, \mathcal{S}]$, the comparison may be described more simply. In this case, we know that $B_L\mathcal{E}$ is the topos $[\mathrm{Fam}_f(\mathbb{C}), \mathcal{S}]$, where $\mathrm{Fam}_f(\mathbb{C})$ as usual denotes the free finite-coproduct-completion of $\mathbb{C}$ (cf. 4.4.17), and $M\mathcal{E}$ is $[\mathrm{CoCart}(\mathbb{C}), \mathcal{S}]$. Since the full embedding $\mathbb{C} \to \mathrm{CoCart}(\mathbb{C})$ obviously factors through $\mathrm{Fam}_f(\mathbb{C})$, we obtain a corresponding factorization

$$\mathcal{E} \longrightarrow B_L\mathcal{E} \longrightarrow M\mathcal{E}$$

of the unit $\delta_{\mathcal{E}}\colon \mathcal{E} \to M\mathcal{E}$, where both factors are essential inclusions. From the universal property of the lower bagdomain, it is easy to verify that the functor B_L preserves inverters in $\mathfrak{BTop}/\mathcal{S}$; hence, by taking a presentation of an arbitrary bounded $\mathcal{S}$-topos $\mathcal{E}$ as an inverter of a 2-cell between diagram toposes (cf. 4.5.1), we may deduce the existence of the corresponding factorization of $\delta_{\mathcal{E}}$, for an arbitrary $\mathcal{E}$. In particular, we deduce that the natural map $B_L\mathcal{E} \to M\mathcal{E}$ is an inclusion for any $\mathcal{E}$. In fact it is an essential inclusion: the left adjoint of its inverse image may be obtained by composing the left adjoint of the inverse image of the inclusion $B_L\mathcal{E} \to MB_L\mathcal{E}$ with the inverse image of $M(\mathcal{E} \to B_L\mathcal{E})$.

But there is another way in which we may factor $\delta_{\mathcal{E}}$. We noted in 4.4.15 that $B_L\mathcal{E}$ is a representing object for pre-geometric morphisms with codomain $\mathcal{E}$; that is, those morphisms which we get by forgetting the requirement that an inverse image functor must preserve the terminal object, but remembering that it must preserve pullbacks. Similarly, $M\mathcal{E}$ represents those $\mathcal{S}$-indexed adjunctions with codomain $\mathcal{E}$ which we get by forgetting altogether the requirement that an inverse image functor should preserve finite limits. If we forget that it should preserve pullbacks, but remember that it should preserve the terminal object, we arrive at the notion of a *probability measure* on $\mathcal{E}$, that is one for which 'the total measure of $\mathcal{E}$ is unity'. This too corresponds to a KZ-monad on $\mathfrak{BTop}/\mathcal{S}$; we may obtain its functor part as follows. Let $P\mathcal{E}$ be the subtopos of $M\mathcal{E}$ corresponding to

the smallest local operator on $M\mathcal{E}$ whose associated sheaf functor maps $\delta_!(1_{\mathcal{E}}) \to 1_{M\mathcal{E}}$ to an isomorphism (cf. A4.5.14(c)). Since the unit map $1_{\mathcal{E}} \to \delta^*\delta_!(1_{\mathcal{E}})$ is an isomorphism, it is clear that δ itself factors through $P\mathcal{E} \to M\mathcal{E}$. More generally, a geometric morphism $f: \mathcal{F} \to M\mathcal{E}$ over $\mathcal{S}$ factors through $P\mathcal{E}$ iff the cocontinuous $\mathcal{S}$-indexed functor $\mathbb{E} \to \mathbb{F}$ to which it corresponds (which is, of course, just the canonical $\mathcal{S}$-indexing of the composite $f^*\delta_!$) preserves 1, i.e. it is an $\mathcal{F}$-valued probability measure on $\mathcal{E}$.

If $\mathcal{E} = [\mathbb{C}, \mathcal{S}]$ for an internal category $\mathbb{C}$, then $P\mathcal{E}$ may be identified with $[\mathrm{Conn}_f(\mathbb{C}), \mathcal{S}]$, where $\mathrm{Conn}_f(\mathbb{C})$ denotes the free completion of $\mathbb{C}$ with respect to connected finite colimits. To see this, we do not need to give an explicit construction of $\mathrm{Conn}_f(\mathbb{C})$; but we note that, as usual, it may be identified with the full subcategory of $[\mathbb{C}^{\mathrm{op}}, \mathbb{S}]$ whose objects are connected finite colimits of representables, from which it is easy to see that the full embedding $\mathbb{C} \to \mathrm{Conn}_f(\mathbb{C})$ is an initial functor. Now a cocontinuous $\mathcal{S}$-indexed functor $[\mathbb{C}, \mathbb{S}] \to \mathbb{F}$ preserves 1 (that is, it is a probability measure on $[\mathbb{C}, \mathcal{S}]$) iff the diagram of shape $\mathbb{C}^{\mathrm{op}}$ in $\mathcal{F}$ which is its restriction to representables (cf. 2.5.8) has colimit 1; so by 2.5.12 we deduce that the same is true of its unique connected-finite-limit-preserving extension to a functor $\mathrm{Conn}_f(\mathbb{C})^{\mathrm{op}} \to \mathbb{F}$. But the preservation of finite connected limits by the latter easily implies, as in 3.2.5, that the total category of the corresponding discrete opfibration is weakly filtered; hence by 2.6.5 it is filtered. So it is a $\mathrm{Conn}_f(\mathbb{C})$-torsor, and hence it induces a geometric morphism $\mathcal{F} \to [\mathrm{Conn}_f(\mathbb{C}), \mathcal{S}]$. The argument in the opposite direction is similar. It follows that, for any $\mathcal{E}$, we can construct a presentation of $P\mathcal{E}$ as an inverter in $\mathfrak{BTop}/\mathcal{S}$ from such a presentation of $\mathcal{E}$, in the same way as we did for $M\mathcal{E}$ but replacing the functor $\mathrm{CoCart}(-)$ by $\mathrm{Conn}_f(-)$.

We next claim that, for any $\mathcal{E}$, we have $B_L(P\mathcal{E}) \simeq M\mathcal{E}$. Again, this could be proved by first considering toposes of the form $[\mathbb{C}, \mathcal{S}]$, and noting that we have $\mathrm{Fam}_f(\mathrm{Conn}_f(\mathbb{C})) \simeq \mathrm{CoCart}(\mathbb{C})$, since any finite colimit can be uniquely decomposed as a finite coproduct of connected colimits. But we can derive the same result more simply, by noting that any cocontinuous $\mathcal{S}$-indexed functor $f: \mathbb{E} \to \mathbb{F}$ can be factored as

$$\mathbb{E} \longrightarrow \mathbb{F}/f(1) \longrightarrow \mathbb{F}$$

(just as in A1.2.9, but without the hypothesis that f preserves pullbacks) where the first factor is cocontinuous and preserves the terminal object. Conversely, given an object B of $\mathcal{F}$ and a cocontinuous 1-preserving functor $\mathbb{E} \to \mathbb{F}/B$, we may compose with the forgetful functor to obtain a cocontinuous functor $\mathbb{E} \to \mathbb{F}$. So $\mathcal{F}$-valued measures on $\mathcal{E}$ are essentially the same thing as bags of $\mathcal{F}$-valued probability measures on $\mathcal{E}$; hence the toposes which represent the two notions must be equivalent.

On the other hand, we do not in general have $P(B_L\mathcal{E}) \simeq M\mathcal{E}$, since we do not have $\mathrm{Conn}_f(\mathrm{Fam}_f(\mathbb{C})) \simeq \mathrm{CoCart}(\mathbb{C})$; although any finite colimit can be written as a connected colimit of finite coproducts (for example, as a coequalizer

(iii) $\Rightarrow$ (i): Define $\theta\colon M\mathcal{E} \to \mathcal{E}$ to be $(\delta_{\mathcal{E}})_+(1_{\mathcal{E}})$. Then $\theta\delta_{\mathcal{E}} \cong 1_{\mathcal{E}}$, since δ is an $\mathcal{S}$-essential inclusion; but idempotent 2-cells split in $\mathfrak{BTop}/\mathcal{S}$, so by 1.1.14 and 4.5.8 $\mathcal{E}$ has an M-algebra structure. (In fact, as is clear from the proof of (i) $\Rightarrow$ (iii), θ itself is the algebra structure map, but we do not need to prove this.)

Now if $\mathcal{E}$ and $\mathcal{E}'$ are M-algebras and $h\colon \mathcal{E} \to \mathcal{E}'$ is an algebra homomorphism, then it follows from the definition of a_+ in (i) $\Rightarrow$ (ii) above that we have

$$a^+(hf) = \theta' \cdot \widetilde{a_! f^* h^*} \cong \theta' \cdot Mh \cdot \widetilde{a_! f^*} \cong h\theta \cdot \widetilde{a_! f^*} = h \cdot a_+(f)$$

so h is cocontinuous. Conversely, if h is cocontinuous, then

$$h\theta \cong h \cdot \delta_+(1_{\mathcal{E}}) \cong \delta_+(h) = \theta' \cdot \widetilde{\delta_! h^*} \cong \theta' \cdot Mh,$$

so it is an algebra homomorphism.

The final assertion is simply a restatement of the fact that $M\mathcal{E}$ is the free M-algebra generated by $\mathcal{E}$. □

It is possible to give a direct proof of (i) $\Rightarrow$ (iii) in 4.5.13, without making use of comma objects or Pitts's theorem, simply by observing that the definition of a_+ in the proof of (i) $\Rightarrow$ (ii) has the property that $a^{\#}a_+$ is isomorphic to the identity whenever $a^*a_!$ is isomorphic to the identity (as it must be if a is an inclusion).

It follows from 1.1.18 that the M-algebras may also be characterized as the objects of $\mathfrak{BTop}/\mathcal{S}$ which are linearly $\mathcal{B}$-complete, where $\mathcal{B}$ is the class of morphisms which occur as the top edges of comma squares whose bottom edges are $\mathcal{S}$-essential. We do not have a simple characterization of this class; but we shall see in C3.3.18 that it may be replaced by the (larger) class of locally connected morphisms.

We remark that the P-algebras in $\mathfrak{BTop}/\mathcal{E}$, where P is the 'probability measures' monad, may be similarly characterized as those which are pointwise cocomplete with respect to the class of those $\mathcal{S}$-essential morphisms $a\colon \mathcal{F} \to \mathcal{G}$ for which the left adjoint $a_!$ preserves 1. (If a is not merely $\mathcal{S}$-essential but $\mathcal{G}$-essential, the latter condition is equivalent to saying that a is connected, i.e. that a^* is full and faithful, as we shall see in C3.3.3. However, in general it is a weaker condition than connectedness.) There is also a generally similar characterization of B_L-algebras as those which are cocomplete with respect to the class of local homeomorphisms (that is, morphisms of the form $\mathcal{G}/A \to \mathcal{G}$ for some object A of $\mathcal{G}$); however, in this case we have to repace the pointwise condition by the linear one, since the class of local homeomorphisms is not stable under 'transfer across comma squares'.

Suggestions for further reading: Bunge [192], Bunge & Carboni [193], Bunge & Funk [197, 198].

of morphisms between coproducts), there are in general many different such representations of a given colimit.

Summarizing the above discussion, we have

Theorem 4.5.12 *For any bounded $\mathcal{S}$-topos $\mathcal{E}$, we have a commutative square of essential inclusions*

$$\begin{tikzcd} \mathcal{E} \arrow[r] \arrow[d] & B_L\mathcal{E} \arrow[d] \\ P\mathcal{E} \arrow[r] & M\mathcal{E} \end{tikzcd}$$

whose diagonal is the unit map $\delta_{\mathcal{E}} \colon \mathcal{E} \to M\mathcal{E}$, and which is natural in $\mathcal{E}$. Moreover, the functors B_L and P both carry KZ-monad structures for which the natural inclusions into M are morphisms of monads; and the composite B_LP is naturally equivalent to M. □

Finally in this section, we characterize the M-algebras in $\mathfrak{BTop}/\mathcal{S}$ in terms of the intrinsic cocompleteness notions introduced in 1.1.16.

Theorem 4.5.13 *Let $\mathcal{S}$ be a topos with a natural number object, and let $\mathcal{A}$ denote the class of $\mathcal{S}$-essential morphisms in $\mathfrak{BTop}/\mathcal{S}$. For a bounded $\mathcal{S}$-topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ admits an M-algebra structure.*

(ii) *$\mathcal{E}$ is pointwise $\mathcal{A}$-cocomplete.*

(iii) *$\mathcal{E}$ is an $\mathcal{A}$-cocomplete $\mathcal{M}$-injective, where $\mathcal{M}$ is the class of inclusions in $\mathcal{A}$.*

Moreover, a geometric morphism between two such toposes is $\mathcal{A}$-cocontinuous iff it is an M-algebra homomorphism; and, for any $\mathcal{E}$, $M\mathcal{E}$ is the free pointwise $\mathcal{A}$-cocompletion of $\mathcal{E}$, in the sense that any geometric morphism $f \colon \mathcal{E} \to \mathcal{F}$ where $\mathcal{F}$ is pointwise $\mathcal{A}$-cocomplete admits a unique (up to isomorphism) $\mathcal{A}$-cocontinuous extension to a morphism $M\mathcal{E} \to \mathcal{F}$.

Proof (i) $\Rightarrow$ (ii): Suppose $(\mathcal{E}, \theta)$ is an M-algebra. Given $a \colon \mathcal{F} \to \mathcal{G}$ in $\mathcal{A}$ and an arbitrary geometric morphism $f \colon \mathcal{F} \to \mathcal{E}$, the composite $a_! f^*$ is a cocontinuous $\mathcal{S}$-indexed functor $\mathbb{E} \to \mathbb{G}$, and so corresponds to a geometric morphism $\mathcal{G} \to M\mathcal{E}$ over $\mathcal{S}$, which we denote by $\widetilde{a_! f^*}$. We define $a_+(f)$ to be the composite $\theta \cdot \widetilde{a_! f^*}$. To verify the adjunction, suppose given $g \colon \mathcal{G} \to \mathcal{E}$; then geometric transformations $a_+(f) \to g$ correspond to geometric transformations $\widetilde{a_! f^*} \to \delta_{\mathcal{E}} g$, since θ is left adjoint to $\delta_{\mathcal{E}}$ by 1.1.13; but these in turn correspond to $\mathcal{S}$-indexed natural transformations $a_! f^* \to g^*$, and hence to transformations $f^* \to a^* g^*$, i.e. to geometric transformations $f \to a^{\#} g$. The Beck–Chevalley condition for comma squares is immediate from 4.1.8.

(ii) $\Rightarrow$ (iii) is immediate from 1.1.19, since we have already observed that inclusions are fully monic in $\mathfrak{Top}$ (and in $\mathfrak{BTop}/\mathcal{S}$).

PART C

TOPOSES AS SPACES

C1

SHEAVES ON A LOCALE

C1.1 Frames and nuclei

In this chapter, our main aim is to describe the 'geometric' origins of topos theory, by setting up the topos of sheaves on a topological space, and indicating how topological properties of spaces and continuous maps may be translated into properties of toposes of sheaves and geometric morphisms between them. However, there is a significant 'generalization' of the notion of topological space, namely the notion of locale, which is of importance in the topos-theoretic context, and almost everything that can be said about toposes of sheaves on spaces carries over without extra effort to toposes of sheaves on locales. Since the notion of locale may not be familiar to the reader, we therefore begin with a couple of sections setting up the category of locales, and describing its relation to the category of spaces.

If X is a topological space, we shall write $\mathcal{O}(X)$ for the lattice of open subsets of X; this lattice is complete (since arbitrary unions of open sets are open), and satisfies the *infinite distributive law*

$$U \wedge \bigvee_{i \in I} V_i = \bigvee_{i \in I} (U \wedge V_i)$$

for all U and all families $(V_i \mid i \in I)$ of elements of $\mathcal{O}(X)$. (This is because the operations of arbitrary join and finite meet in $\mathcal{O}(X)$ coincide with the set-theoretic operations of union and intersection, which are well known to satisfy all possible distributive laws. However, the lattice-theoretic dual of this law may fail in $\mathcal{O}(X)$, since the meet of an infinite family of open sets is not their intersection but the interior of their intersection.) Moreover, if $f: X \to Y$ is a continuous map of topological spaces, the mapping $f^{-1}: \mathcal{O}(Y) \to \mathcal{O}(X)$ preserves finite meets and arbitrary joins. These two observations motivate the following definition:

Definition 1.1.1 A *frame* is a complete lattice A satisfying the infinite distributive law

$$a \wedge \bigvee_{i \in I} b_i = \bigvee_{i \in I} (a \wedge b_i)$$

for all $a \in A$ and all families $(b_i \mid i \in I)$ of elements of A. A *frame homomorphism* $h: A \to B$ is a mapping preserving finite meets and arbitrary joins. We write **Frm** for the category of frames and frame homomorphisms.

Remark 1.1.2 A complete lattice is a frame iff it is a Heyting algebra (cf. A1.5.11), since the Adjoint Functor Theorem tells us that $a \wedge (-): A \to A$ preserves arbitrary joins iff it has a right adjoint $a \Rightarrow (-)$. (This is, of course, a special case of the assertion that geometric categories are Heyting categories, which we noted in A1.4.18, since a frame is just a geometric category which is a poset.) However, frame homomorphisms do not in general preserve the implication operation.

In particular, we note that any complete Boolean algebra is a frame. Since frame homomorphisms are in particular lattice homomorphisms, they preserve complements of elements when they exist; hence a frame homomorphism between complete Boolean algebras is the same thing as a complete Boolean homomorphism (i.e. a map preserving arbitrary meets and joins). We shall have more to say about the relationship between frames and complete Boolean algebras in 1.1.21 below.

In this section we study the 'algebraic' aspects of frame theory; its 'topological' aspects will be the subject of the next section. From the algebraic point of view, one of the most important results about frames is that free frames exist, i.e. that the forgetful functor **Frm** $\to$ **Set** has a left adjoint. There are several different constructions of this left adjoint, which correspond to different factorizations of the forgetful functor through various 'intermediate' categories of posets; we shall want to make use of all of them at different times in what follows, and so we devote the next few paragraphs to setting them up.

We shall wish to distinguish notationally between the categories **jSLat** and **mSLat** of join-semilattices and meet-semilattices; abstractly, these are of course the same category, but their forgetful functors to **Poset** are different. (Incidentally, for us a semilattice will always have a unit element: that is, a join-semilattice has a bottom element and a meet-semilattice has a top element.) **DLat** will denote the category of (finitary) distributive lattices, and **CjSLat** the category of complete join-semilattices: that is, an object of **CjSLat** is a poset with arbitrary joins (and hence also arbitrary meets), and a morphism of **CjSLat** is a function preserving arbitrary joins (equivalently, having a right adjoint), but not necessarily preserving meets. We shall use the term *domain* for a poset having joins of (upwards) directed subsets, and write **Dom** for the category of domains (and functions preserving directed joins). Finally, a *preframe* is defined to be a poset with finite meets and directed joins, which satisfies the infinite distributive law of 1.1.1 whenever the family $(b_i \mid i \in I)$ is directed (though the law may fail for other families, even if the joins happen to exist), and a *preframe homomorphism* is a function preserving finite meets and directed joins: we write **PFrm** for the category of preframes.

Clearly, we have forgetful functors

$$\begin{tikzcd}
\mathbf{Frm} \arrow[rr] \arrow[dr] \arrow[dd] & & \mathbf{CjSLat} \arrow[dr] \arrow[dd] & . \\
& \mathbf{PFrm} \arrow[rr] \arrow[dd] & & \mathbf{Dom} \arrow[dd] \\
\mathbf{DLat} \arrow[rr] \arrow[dr] & & \mathbf{jSLat} \arrow[dr] & \\
& \mathbf{mSLat} \arrow[rr] & & \mathbf{Set}
\end{tikzcd}$$

Lemma 1.1.3 *All the functors in the above diagram have left adjoints.*

Proof We shall not give explicit constructions of all the left adjoints (those that we do not mention can easily be shown to exist using the Adjoint Functor Theorem, or by the adjoint lifting theorem A1.1.3). However, we shall give descriptions of the ones of particular interest to us.

First, the free join-semilattice on a set A is the set KA of finite subsets of sA, ordered by inclusion (cf. D5.4.9): given a function $f\colon A \to B$ where B is a join-semilattice, its unique extension to a join-preserving map $\overline{f}\colon KA \to B$ is given by $\overline{f}(S) = \bigvee\{f(s) \mid s \in S\}$. Hence also the free meet-semilattice on A is simply KA^{op}. A similar argument shows that the free complete join-semilattice on a set A is the set PA of all subsets of A, ordered by inclusion. And the free domain on a set A is just A itself, with the discrete ordering; note that this is a domain, since its only directed subsets are singletons.

The free functor $\mathbf{mSLat} \to \mathbf{Frm}$ sends a meet-semilattice A to the set LA of all lower (i.e. downwards-closed) subsets of A, ordered by inclusion; the unit map $\eta\colon A \to LA$ sends a to the principal ideal $\downarrow(a) = \{b \in A \mid b \leq a\}$. It is clear that LA is a frame (and indeed a coframe), since it is closed under arbitrary meets and joins in PA; and η preserves all finite meets (indeed, all meets which happen to exist in A). Now if B is a frame and $f\colon A \to B$ is a map preserving finite meets, the only possible candidate for a frame homomorphism $\overline{f}\colon LA \to B$ extending f is given by

$$\overline{f}(S) = \bigvee\{f(s) \mid s \in S\},$$

since each lower set is the union of the principal ideals which it contains; this map clearly does preserve joins (and extend f), and it also preserves finite meets since $\overline{f}(A) = f(1_A) = 1_B$ and

$$\begin{aligned}\overline{f}(S) \wedge \overline{f}(T) &= \bigvee\{f(s) \wedge f(t) \mid s \in S, t \in T\} \\ &= \bigvee\{f(s \wedge t) \mid s \in S, t \in T\} \\ &= \bigvee\{f(u) \mid u \in S \cap T\} = \overline{f}(S \cap T)\end{aligned}$$

where we used the infinite distributive law in B at the first step.

A similar argument shows that the free distributive lattice on a meet-semilattice A is the set of finitely-generated lower sets in A (that is, sets which are finite unions of principal ideals); and the free preframe on A is the set IA of *ideals* (= directed lower sets) in A. (Note that the intersection of two ideals S and T in a meet-semilattice is an ideal; for if a and b are two elements of $S \cap T$, we can find upper bounds c and d for $\{a, b\}$ in S and T respectively, and then $c \wedge d$ is an upper bound for $\{a, b\}$ in $S \cap T$.)

For a *join*-semilattice A, the construction of IA yields the free complete join-semilattice on A: it is not necessarily closed under finite unions (or intersections) in PA, but it does have binary joins given by

$$S \vee T = \{s \vee t \mid s \in S, t \in T\}$$

(and the mapping $a \mapsto \downarrow(a)$ preserves them), as well as a least element $\{0\}$. Since it also has directed joins (which are simply unions), it is a complete join-semilattice: the proof that it is freely generated by A is similar to those already given. The same construction, applied to a distributive lattice A, yields a frame which is (unsurprisingly) the free frame generated by A.

We shall not give explicit constructions of the free functors **CjSLat** $\to$ **Frm** and **PFrm** $\to$ **Frm** here, since we shall not need them; but we shall discuss the relationship between **Frm** and these two categories further in 1.1.10 below. □

In passing, we note

Scholium 1.1.4 *The free frame on any set A is of the form $\mathcal{O}(X)$ for a suitable topological space X.*

Proof From 1.1.3, we see that the free frame on A may be constructed as $L(KA^{\mathrm{op}})$, which is manifestly a topology on the set KA. Alternatively, we may describe it as the open-set lattice of the A-fold cartesian power of the *Sierpiński space* $S = \{0, 1\}$ with topology $\mathcal{O}(S) = \{\emptyset, \{1\}, S\}$. To establish the isomorphism between $\mathcal{O}(S^A)$ and $L(KA^{\mathrm{op}})$, note that a nonempty 'open rectangle' $\prod_{a \in A} U_a$ in S^A is determined by the set $\{a \in A \mid U_a = \{1\}\}$, which is necessarily finite (but the inclusion ordering on open rectangles is the reverse of the inclusion ordering on finite subsets of A), and that an arbitrary open subset of S^A is

determined by the set of open rectangles which it contains, by the definition of the product topology. □

Since the functor $\mathcal{O}\colon \mathbf{Sp}^{\mathrm{op}} \to \mathbf{Set}$, which sends a space to its set of open subsets, is representable by the Sierpiński space, the result just established says that elements of the free frame on A correspond (to continuous maps $S^A \to S$, and hence) to natural transformations $\mathcal{O}^A \to \mathcal{O}$. In other words, the theory of frames is exactly the (infinitary) algebraic theory of the functor $\mathcal{O}$, in the sense of [724].

All the categories in the diagram before 1.1.3, except for **Dom** and **PFrm**, are equationally presented (that is, they are the categories of models of (infinitary) algebraic theories, cf. D1.1.7(a)): hence they are all monadic over **Set**, and in particular they are complete and cocomplete. (The two exceptions are *not* monadic over **Set**: the monad induced by the forgetful functor **Dom** $\to$ **Set** and its left adjoint is the trivial one whose functor part is the identity, and that induced by the forgetful functor **PFrm** $\to$ **Set** is the same as that induced by **mSLat** $\to$ **Set**, since it is easily seen that every ideal in a meet-semilattice of the form KA^{op} is principal. On the other hand, the forgetful functor **PFrm** $\to$ **mSLat** is monadic; and **PFrm** is still complete and cocomplete. These results are proved in [548], but we shall not give the details here.)

Proposition 1.1.5 *The categories* **CjSLat** *and* **PFrm** *both have (symmetric) closed monoidal structures.*

Proof We consider first the category **CjSLat**. The algebraic theory of complete join-semilattices is commutative (also called entropic, or autonomous in the sense of [725]): that is, each of its operations commutes with each of the others (and with itself). Hence, given two complete semilattices A and B, the set of **CjSLat**-morphisms $A \to B$ is closed under (pointwise) joins in the set of all functions $A \to B$, and so forms a complete semilattice $[A, B]$. If C is a third object of **CjSLat**, then **CjSLat** morphisms $C \to [A, B]$ are easily seen to correspond to functions $f\colon C \times A \to B$ which are 'bihomomorphisms' in the sense that $c \mapsto f(c, a)$ is a homomorphism for each fixed $a \in A$ and $a \mapsto f(c, a)$ is a homomorphism for each fixed $c \in C$. So we define $A \otimes B$ to be the codomain of the universal bihomomorphism with domain $A \times B$: explicitly, it is freely generated by symbols $a \otimes b$ ($a \in A$, $b \in B$), subject to the relations

$$\bigvee_{i \in I}(a_i \otimes b) = (\bigvee_{i \in I} a_i) \otimes b \quad \text{and} \quad \bigvee_{j \in J}(a \otimes b_j) = a \otimes (\bigvee_{j \in J} b_j)$$

for all a, b, a_i and b_j. Just as for the usual tensor product of abelian groups (or of modules over a commutative ring), it is straightforward to verify that this construction is commutative and associative up to coherent isomorphisms (for example, the 'twist' isomorphism $A \otimes B \to B \otimes A$ is the unique morphism sending $a \otimes b$ to $b \otimes a$ for each (a, b)), that the free complete semilattice on one generator is a unit for it (again, up to coherent isomorphism), and that $(-) \otimes A$ is left

adjoint to $[A, -]$. (The free complete semilattice on one generator is simply the power-set $P1$, by the results of 1.1.3; we shall denote it by Ω, because if we replace **Set** by a more general topos $\mathcal{E}$ in this discussion, the unit becomes the subobject classifier of $\mathcal{E}$. Note also that the generator of Ω is its top element $\top$.)

The argument for **PFrm** is similar: if A is a preframe, the infinite distributive law ensures that the binary meet operation $A \times A \to A$ preserves directed joins, since if $(a_i \mid i \in I)$ and $(b_i \mid i \in I)$ are directed families (indexed by the same directed set I) then $(a_i \wedge b_i \mid i \in I)$ is final in the family $(a_i \wedge b_j \mid i, j \in I)$. Thus the theory of preframes is 'commutative' in an appropriate sense: in particular, if A and B are preframes then the set $[A, B]$ of preframe homomorphisms $A \to B$ is closed under (pointwise) finite meets and directed joins in the set of all maps $A \to B$, and hence forms a preframe. As before, the tensor product of two preframes A and B (which we shall denote by $A \odot B$, since we wish to distinguish it from the **CjSLat** tensor product) is the codomain of the universal 'bihomomorphism' from $A \times B$: equivalently, it is generated by elements $a \odot b$ ($a \in A$, $b \in B$) subject to relations which ensure that $a \mapsto a \odot b$ is a preframe homomorphism for each b, and $b \mapsto a \odot b$ is a homomorphism for each a. The unit of the tensor product is the free preframe on one generator: by the results of 1.1.3, this is the preframe $I(K1^{\mathrm{op}})$, but since $K1^{\mathrm{op}}$ is isomorphic to $K1$ (even constructively, cf. D5.4.5(ii)) we may identify this with $I(K1) \cong P1 = \Omega$. (However, the generator of Ω as a preframe is its *bottom* element $\bot$.) □

Remark 1.1.6 Although **Frm** itself does not have a closed structure, its hom-sets do have an obvious partial ordering, given by $f \leq g$ iff $f(a) \leq g(a)$ for all a in the domain of f; and this partial ordering has (pointwise) directed joins, since a directed join of frame homomorphisms is both a preframe homomorphism and a complete-join-semilattice homomorphism. Thus **Frm** is enriched over the category **Dom** – a fact which we shall subsequently use when working in the opposite category **Loc**.

Remark 1.1.7 In fact **CjSLat** has even more categorical structure than that described in 1.1.5. It has a self-duality (i.e. a functor $(-)^* : \mathbf{CjSLat} \to \mathbf{CjSLat}^{\mathrm{op}}$ whose square is the identity): for an object A, A^* is simply the opposite poset A^{op} (we have already noted that this has arbitrary joins if A does), and for a morphism $f : A \to B$, f^* is the right adjoint of f regarded as a morphism $B^{\mathrm{op}} \to A^{\mathrm{op}}$. Moreover, the duality is enriched in the sense that we have $[A, B] \cong [B^*, A^*]$, naturally in A and B: for the pointwise ordering on the right adjoints of morphisms $A \to B$ is the opposite of that on the original morphisms, and it becomes reversed again when we dualize the objects A and B. It follows in particular that we have natural isomorphisms $A^* \cong [\Omega, A^*] \cong [A, \Omega^*]$, i.e. the duality is representable in the enriched sense. (Of course, Ω^* is classically isomorphic to Ω, but they are different in a non-Boolean topos, and so we choose to distinguish between them here.) Hence also we have natural isomorphisms

$$[A, B] \cong [B^*, A^*] \cong [B^*, [A, \Omega^*]] \cong [(B^* \otimes A), \Omega^*] \cong (B^* \otimes A)^* \cong (A \otimes B^*)^*$$

or equivalently $(A \otimes B) \cong [A, B^*]^*$: the interpretation of this last isomorphism is that we may identify $A \otimes B$ with the poset of *Galois connections* (that is, contravariant adjunctions on the right) between A and B. Summarizing the foregoing, we say that **CjSLat** is a $*$-*autonomous category* in the sense of [72].

We recall that, given a symmetric monoidal category $(\mathcal{C}, \otimes, I)$, we may form the category $\mathbf{CMon}(\mathcal{C})$ of commutative monoids in $\mathcal{C}$, i.e. objects A equipped with morphisms $m\colon A \otimes A \to A$ and $e\colon I \to A$ satisfying the appropriate diagrammatic form of the equations for commutative monoids. We also recall the following result (cf. [368]):

Lemma 1.1.8 *Let $(\mathcal{C}, \otimes, I)$ be a symmetric monoidal category. Then*

(i) *The category $\mathbf{CMon}(\mathcal{C})$ has finite coproducts; and the forgetful functor $\mathbf{CMon}(\mathcal{C}) \to \mathcal{C}$ is universal among functors from categories $\mathcal{D}$ with finite coproducts to $\mathcal{C}$ which map the cocartesian monoidal structure on $\mathcal{D}$ to the given monoidal structure on $\mathcal{C}$.*

(ii) *If the monoidal structure on $\mathcal{C}$ is closed, then the forgetful functor $\mathbf{CMon}(\mathcal{C}) \to \mathcal{C}$ creates any filtered colimits which exist in $\mathcal{C}$.*

Proof (i) It is straightforward to verify that if (A, m, e) and (A', m', e') are commutative monoids in $\mathcal{C}$, then $A \otimes A'$, with its obvious monoid structure, is their coproduct in $\mathbf{CMon}(\mathcal{C})$: the coprojections are given by $1 \otimes e'\colon A \cong A \otimes I \to A \otimes A'$ and $e \otimes 1\colon A' \cong I \otimes A' \to A \otimes A'$, and if $f\colon A \to A''$ and $f'\colon A' \to A''$ are two monoid homomorphisms the induced homomorphism $A \otimes A' \to A''$ is $m''(f \otimes f')$. Similarly, the unit I with its obvious monoid structure is an initial object of $\mathbf{CMon}(\mathcal{C})$. On the other hand, if $\mathcal{D}$ is any category with finite coproducts, then each object B of $\mathcal{D}$ has a unique commutative monoid structure given by the codiagonal map $B + B \to B$ and the unique morphism $0 \to B$; so any functor $\mathcal{D} \to \mathcal{C}$ which maps the finite coproduct structure onto $\otimes$ has a unique factorization through $\mathbf{CMon}(\mathcal{C}) \to \mathcal{C}$.

(ii) Suppose $A = \varinjlim_{i \in I} A_i$ is a filtered colimit in $\mathcal{C}$, where each A_i has a commutative monoid structure and the transition maps $A_i \to A_j$ are monoid homomorphisms. Since $A \otimes (-)$ preserves colimits, it is easy to see that we have $A \otimes A \cong \varinjlim_{i,j \in I \times I}(A_i \otimes A_j)$; but the fact that I is filtered implies that the diagonal $I \to I \times I$ is final (cf. B2.5.12), and so $A \otimes A \cong \varinjlim_{i \in I}(A_i \otimes A_i)$. Hence the multiplications $m_i\colon A_i \otimes A_i \to A_i$ induce a unique morphism $m\colon A \otimes A \to A$; and we similarly have a morphism $e\colon I \to A$ induced by any of the $e_i\colon I \to A_i$ (since I is nonempty and connected). The verification that these morphisms are a commutative monoid structure on A, and that they make A into the colimit of the A_i in $\mathbf{CMon}(\mathcal{C})$, is straightforward. $\square$

We note that the argument of 1.1.8(ii) also works for reflexive coequalizers (cf. A1.2.12); thus we could have replaced the word 'filtered' by 'sifted' in the

statement. Also, since arbitrary colimits may be constructed from finite coproducts, reflexive coequalizers and filtered colimits, 1.1.8 in principle tells us how to construct arbitrary colimits in $\mathbf{CMon}(\mathcal{C})$ in terms of $\mathcal{C}$.

More explicitly, the identification of finite coproducts in $\mathbf{CMon}(\mathcal{C})$ with tensor products may be extended to pushouts, by the following device. Given a commutative monoid (A, m, e) in $\mathcal{C}$, we may define an *A-module* to be an object B of $\mathcal{C}$ equipped with an action map $\beta\colon A \otimes B \to B$ which is associative and unitary in the obvious sense. Provided $\mathcal{C}$ has reflexive coequalizers, and the functor $A \otimes (-)$ preserves them, the category A-$\mathbf{Mod}(\mathcal{C})$ of A-modules in $\mathcal{C}$ has a monoidal structure, with the tensor product $B \otimes_A C$ of two modules defined to be the coequalizer of

$$B \otimes A \otimes C \underset{1 \otimes \gamma}{\overset{\beta \otimes 1}{\rightrightarrows}} B \otimes C$$

(note that since A is commutative we do not have to distinguish between left and right A-modules), and unit given by A itself. Now if $f\colon A \to B$ is a morphism of $\mathbf{CMon}(\mathcal{C})$, then B acquires an A-module structure given by

$$A \otimes B \xrightarrow{f \otimes 1} B \otimes B \xrightarrow{m} B;$$

moreover, the multiplication of B factors through the quotient map $B \otimes B \twoheadrightarrow B \otimes_A B$, making B into a commutative monoid in A-$\mathbf{Mod}(\mathcal{C})$. Conversely, any commutative monoid B in A-$\mathbf{Mod}(\mathcal{C})$ is in particular a commutative monoid in $\mathcal{C}$, and its unit map in the former category becomes a monoid homomorphism $A \to B$ in the latter. These two constructions establish an equivalence between the co-slice category $A\backslash\mathbf{CMon}(\mathcal{C})$ and $\mathbf{CMon}(A\text{-}\mathbf{Mod}(\mathcal{C}))$; so, in particular, we obtain

Corollary 1.1.9 *Suppose $\mathcal{C}$ is a symmetric monoidal category with reflexive coequalizers which are preserved by the functors $A \otimes (-)$, and let $f\colon A \to B$ and $g\colon A \to C$ be two commutative monoid homomorphisms in $\mathcal{C}$ with common domain. Then (the codomain of) their pushout may be identified with the A-module tensor product $B \otimes_A C$.* □

The significance of 1.1.8 and 1.1.9 for us is contained in

Proposition 1.1.10 *The category* **Frm** *may be identified with coreflective (full) subcategories of both* $\mathbf{CMon}(\mathbf{CjSLat})$ *and* $\mathbf{CMon}(\mathbf{PFrm})$.

Proof If A is a frame, then the infinite distributive law of 1.1.1 says precisely that the binary meet map $\wedge\colon A \times A \to A$ is a bihomomorphism of complete join-semilattices; so it induces a map $m\colon A \otimes A \to A$. And the top element 1 of A corresponds to a unique **CjSLat** morphism $e\colon \Omega \to A$. The fact that m and e satisfy the equations for commutative monoids is immediate from the

fact that $\wedge$ and 1 do so (with respect to the cartesian monoidal structure); and it is clear that if B is another frame then a **CjSLat** morphism $A \to B$ is a morphism of commutative monoids iff it also preserves finite meets, i.e. iff it is a frame homomorphism. Thus we have identified **Frm** with a full subcategory of **CMon**(**CjSLat**).

Of course, not every commutative monoid in **CjSLat** is of this form; but we may characterize those which are, as follows: (A, m, e) is a frame iff $a \leq e(\top)$ and $a \leq m(a \otimes a)$ for all $a \in A$. For these conditions clearly hold in a frame; conversely, if the conditions hold, then for any $a, b \in A$ we have $m(a \otimes b) \leq m(a \otimes e(\top)) = a$ and $m(a \otimes b) \leq b$ similarly, and if $c \leq a$ and $c \leq b$ then $c \leq m(c \otimes c) \leq m(a \otimes b)$, so $m(a \otimes b)$ is the meet of a and b. Now, if (A, m, e) is an arbitrary commutative monoid in **CjSLat**, consider the set

$$A_f = \{a \in A \mid a \leq e(\top) \text{ and } a \leq m(a \otimes a)\};$$

it is clear that A_f is a sub-complete-semilattice of A, since if a is the join of a family $(a_i \mid i \in I)$ then $m(a \otimes a) \geq m(a_i \otimes a_i) \geq a_i$ for each i. Moreover, if a and b are in A_f so is $m(a \otimes b)$, using the fact that m is commutative and associative; and hence m maps $A_f \otimes A_f$ into A_f. Similarly, since $e(\top) \in A_f$, e factors through $A_f \rightarrowtail A$, and thus A_f is a sub-commutative-monoid of A. By the remarks above, it is a frame; and it is also clear that any commutative monoid homomorphism from a frame to A must take values in A_f. So we have constructed the required coreflection.

The arguments in the preframe case are very similar. This time we consider the binary join map $\vee \colon A \times A \to A$ of a frame A as a preframe bihomomorphism: it is a bihomomorphism for finite meets by the (finitary) distributive law, and for directed joins (in fact for all nonempty joins) since $a \vee (\bigvee_{i \in I} b_i) = \bigvee_{i \in I}(a \vee b_i)$ provided a appears at least once on the right-hand side. The induced preframe morphism $m \colon A \odot A \to A$, together with the unique preframe morphism $e \colon \Omega \to A$ sending the preframe generator $\perp$ of Ω to 0_A, defines a commutative monoid structure on A; and a preframe homomorphism between frames is a homomorphism of commutative monoids iff it preserves finite joins, i.e. iff it is a frame homomorphism. A commutative monoid (A, m, e) in **PFrm** arises in this way from a frame iff $e(\perp)$ is the least element of A and $m(a \odot a) \leq a$ for all $a \in A$; using this, we may prove as before that **Frm** is coreflective in **CMon**(**PFrm**).

$\square$

Remark 1.1.11 It follows from 1.1.8(i) and 1.1.10 that if A and B are frames then their coproduct in **Frm** may be identified with both their complete-semilattice tensor product $A \otimes B$ and their preframe tensor product $A \odot B$; so these two are isomorphic as posets. However, the isomorphism does not identify the generators of $A \otimes B$ with those of $A \odot B$; if we write ν_1 and ν_2 for the coprojections in **Frm**, then $a \otimes b$ is $\nu_1(a) \wedge \nu_2(b)$, and $a \odot b$ is $\nu_1(a) \vee \nu_2(b)$. It

follows easily that they are related to each other by the equations

$$(a \odot b) = (a \otimes 1_B) \vee (1_A \otimes b) \quad \text{and} \quad (a \otimes b) = (a \odot 0_B) \wedge (0_A \odot b)\,.$$

Similar remarks apply to pushouts in **Frm**, using 1.1.9.

The representations of frames as commutative monoids in **CjSLat** and in **PFrm** will be extensively exploited in Sections C3.1 and C3.2 respectively. However, before leaving the topic for now, we note one useful application of 1.1.10:

Corollary 1.1.12 *Let $(A_i \mid i \in I)$ be the vertices of a filtered diagram in* **Frm**, *whose transition maps $A_i \to A_j$ are all injective. Then the legs $A_i \to \varinjlim_{i\in I} A_i$ of the colimiting cone in* **Frm** *are also injective.*

Proof By 1.1.8(ii) and 1.1.10, it suffices to prove the corresponding result in the category **CjSLat**. But in the latter category we can appeal to the duality of 1.1.7: in other words, it suffices to show that, given a cofiltered diagram of complete semilattices $(B_i \mid i \in I)$ and surjective homomorphisms, the induced maps $\lambda_j \colon \varprojlim_{i\in I} B_i \to B_j$ are surjective. Let $b \in B_i$ for some i; since the forgetful functor **CjSLat** $\to$ **Set** creates limits, we have to construct a family of elements $(b_j \in B_j \mid j \in I)$ such that $b_i = b$ and such that $\beta(b_k) = b_j$ for each edge $\beta \colon B_k \to B_j$ of the diagram. For each j, we may choose morphisms $\alpha \colon B_k \to B_i$ and $\gamma \colon B_k \to B_j$ in the diagram with common domain; we define $b_j = \gamma(\alpha^*(b))$, where α^* is the dual of α. It is clear that, if we can show b_j to be independent of the choice of the pair (α, γ), then $(b_j \mid j \in I)$ will be a compatible family. But if we had made a different choice (α', γ'), we could find $\delta \colon B_l \to B_k$ and $\delta' \colon B_l \to B_{k'}$ in the diagram with $\alpha\delta = \alpha'\delta'$ and $\gamma\delta = \gamma'\delta'$, so that

$$\begin{aligned}
\gamma(\alpha^*(b)) &= \gamma\delta(\delta^*\alpha^*(b)) \\
&= \gamma'\delta'(\delta'^*\alpha'^*(b)) \\
&= \gamma'(\alpha'^*(b))
\end{aligned}$$

where we have used the fact that δ and δ' are surjective and so the counits of the adjunctions $(\delta \dashv \delta^*)$ and $(\delta' \dashv \delta'^*)$ are identities. □

Since **Frm** is monadic over **Set**, it is an effective regular category as defined in Section A1.3: in particular, for any frame A, there is a bijection between regular quotients of A (that is, isomorphism classes of regular epimorphisms $A \twoheadrightarrow B$) and equivalence relations on A. Since monomorphisms in **Frm** are injective frame maps, the latter may be identified with frame congruences on A, that is subsets $R \subseteq A \times A$ which are both equivalence relations on (the underlying set of) A and subframes of $A \times A$. But there are a number of other ways of representing regular quotients of A, which will be of use to us. First we note that, since a frame congruence R is closed under arbitrary joins in $A \times A$, each R-equivalence class has a greatest member (the join of all its members); thus we can represent

R by the mapping (j, say) $A \to A$ which sends each element of A to the greatest member of its R-equivalence class, since we then have $(a, b) \in R$ iff $j(a) = j(b)$. And the properties of j needed to ensure that the relation defined in this way is a frame congruence are simply

(i) $a \leq j(a)$;
(ii) $j(a) = j(j(a))$; and
(iii) $j(a \wedge b) = j(a) \wedge j(b)$

for all $a, b \in A$. We call such a mapping a *nucleus* on A. (The above conditions should be compared with the definition of a local operator in a topos (A4.4.1): a local operator is simply an internal nucleus on the internal frame Ω, and much of the discussion of nuclei which follows in this section exactly parallels the discussion of local operators in Section A4.5.) Further, a nucleus j is determined by its image (equivalently, by (ii), its *fixset* $A_j = \{a \in A \mid j(a) = a\}$), since conditions (i) and (ii) (plus the fact that j is order-preserving) say that j is left adjoint to the inclusion $A_j \to A$. Moreover, in this context condition (iii) on j is equivalent (by A4.3.1) to the assertion that A_j is an exponential ideal as defined in Section A1.5, i.e. that $(a \Rightarrow b) \in A_j$ whenever $b \in A_j$, where $\Rightarrow$ is the Heyting implication in A (cf. 1.1.2). We have thus established

Proposition 1.1.13 *For any frame A, there are bijections between any two of the following sets:*

(i) *the set of regular quotients of A, i.e. isomorphism classes of surjective frame homomorphisms with domain A;*
(ii) *the set of frame congruences on A;*
(iii) *the set of nuclei on A;*
(iv) *the set of fixsets in A, i.e. subsets which are exponential ideals and closed under arbitrary meets.* □

We remark that it is possible to proceed directly from a regular epimorphism $q: A \twoheadrightarrow B$ to the nucleus j on A which corresponds to it: in fact j is the composite q_*q, where $q_*: B \to A$ is the right adjoint of q. (This is because $q_*(b)$ is the largest $a \in A$ satisfying $q(a) = b$.) And, given j, we may recover q directly: A_j, ordered as a subset of A, is isomorphic to the quotient A/R (since it contains one element of each R-equivalence class) and hence a frame; and j itself, considered as a map $A \to A_j$, is a (surjective) frame homomorphism (it preserves joins because it is left adjoint to the inclusion).

Remark 1.1.14 The three conditions in the definition of a nucleus may be combined into a single equation: *for any frame A, a function $j: A \to A$ is a nucleus iff it satisfies the condition that*

$$(a \Rightarrow j(b)) = (j(a) \Rightarrow j(b))$$

for all $a, b \in A$. To prove this, suppose first that j is a nucleus; then from $a \leq j(a)$ we obtain $(a \Rightarrow j(b)) \geq (j(a) \Rightarrow j(b))$, since $\Rightarrow$ is order-reversing in its first variable. But we also have

$$\begin{aligned} j(a) \wedge (a \Rightarrow j(b)) &\leq j(a) \wedge j(a \Rightarrow j(b)) \\ &= j(a \wedge (a \Rightarrow j(b))) \\ &\leq j(j(b)) = j(b), \end{aligned}$$

so $(a \Rightarrow j(b)) \leq (j(a) \Rightarrow j(b))$. Conversely, suppose j satisfies the condition. Putting $a = b$, we obtain $(a \Rightarrow j(a)) = 1_A$, so $a \leq j(a)$ for all a. Similarly, putting $a = j(b)$ yields $j(j(b)) \leq j(b)$, whence $j(j(b)) = j(b)$. Next, if $a \leq b$ then $a \leq j(b)$, whence we obtain $(j(a) \Rightarrow j(b)) = 1_A$, so j is order-preserving. Finally, from $a \wedge b \leq j(a \wedge b)$ we obtain $a \leq (b \Rightarrow j(a \wedge b)) = (j(b) \Rightarrow j(a \wedge b))$, so $a \wedge j(b) \leq j(a \wedge b)$, and a similar argument yields $j(a) \wedge j(b) \leq j(a \wedge b)$; but the reverse inequality follows from the fact that j is order-preserving. However, in practice it is usually easier to verify the three separate conditions, for a given j, than to verify the single equation given here.

Of course, each of the sets described in 1.1.13 has a canonical ordering; but some of the bijections between them are order-reversing. If we order congruences by inclusion, this corresponds to the pointwise ordering on nuclei (i.e. $j \leq k$ iff $j(a) \leq k(a)$ for all $a \in A$), but it is the reverse of the inclusion ordering on fixsets, since larger nuclei have fewer fixed points. For definiteness, we shall deal with nuclei in what follows, and write $N(A)$ for the set of all nuclei on A with its pointwise ordering.

Proposition 1.1.15 *For any frame A, $N(A)$ is a frame.*

Proof It is straightforward to verify that the (pointwise) meet of an arbitrary family of nuclei is a nucleus; this defines arbitrary meets in $N(A)$, and shows that it is a complete lattice. However, joins in $N(A)$ are not in general computed pointwise; so it is not possible to deduce the infinite distributive law in $N(A)$ directly from that in A. One possible proof, which may be found in [520], involves constructing the Heyting implication operation on $N(A)$ and appealing to 1.1.2; we shall give a simpler proof, due to J. T. Wilson [1229], which in fact shows that the fixsets in A form a coframe in the inclusion ordering. For this, we note first that an arbitrary intersection of fixsets is a fixset. The join of two fixsets B and C may be described as the set $\{b \wedge c \mid b \in B, c \in C\}$; this follows from the fact that meets of nuclei are computed pointwise, since an element a of A is $(j \wedge k)$-fixed iff it is the meet of $j(a)$ and $k(a)$. Now suppose B and C_i, $i \in I$, are fixsets, and let $a \in \bigcap_{i \in I}(B \vee C_i)$. Then we have $a = b \wedge c_i$ for each i, where b is the least element of B above a and c_i is the least element of C_i above a. So

$$(b \Rightarrow a) = (b \Rightarrow (b \wedge c_i)) = (b \Rightarrow c_i) \in C_i \quad \text{for each } i$$

since C_i is an exponential ideal, whence $a = (b \wedge a) = (b \wedge (b \Rightarrow a)) \in B \vee \bigcap_{i \in I} C_i$. Thus we have

$$\bigcap_{i \in I}(B \vee C_i) \subseteq B \vee \bigcap_{i \in I} C_i;$$

but the reverse inclusion holds in any complete lattice. □

We next consider some important examples of nuclei and fixsets: the first three of the following examples should be compared with their topos-theoretic analogues introduced in A4.5.1, A4.5.3 and A4.5.9 respectively.

Examples 1.1.16 (*a*) Let A be a frame, a an element of A. The principal ideal $\downarrow(a)$ is easily seen to be a frame, and the infinite distributive law implies that $(-) \wedge a \colon A \to \downarrow(a)$ is a (surjective) frame homomorphism. So by 1.1.13 it corresponds to a nucleus on A. Since the right adjoint of $(-) \wedge a$ is $a \Rightarrow (-)$, it is easy to see that this nucleus is simply $a \Rightarrow (-) \colon A \to A$. We denote this nucleus by $o(a)$, and call it the *open nucleus* determined by A, for reasons which will become clear in 1.2.6 below.

(*b*) Again, let a be an element of a frame A. The principal filter $\uparrow(a)$ generated by a is also a frame, and $(-) \vee a \colon A \to \uparrow(a)$ is a surjective frame homomorphism by distributivity. The right adjoint of this homomorphism is simply the inclusion $\uparrow(a) \to A$, and so the corresponding nucleus is simply $(-) \vee a$ (and the corresponding fixset is $\uparrow(a)$ itself). We denote this nucleus by $c(a)$, and call it the *closed nucleus* determined by a. We note that $o(a)$ and $c(a)$ are complementary elements of $N(A)$, for any a: for their pointwise meet $o(a) \wedge c(a)$ satisfies

$$\begin{aligned}(o(a) \wedge c(a))(b) &= (a \Rightarrow b) \wedge (a \vee b) \\ &= (a \wedge (a \Rightarrow b)) \vee (b \wedge (a \Rightarrow b)) \\ &= (a \wedge b) \vee b = b\end{aligned}$$

using standard identities for Heyting algebras (cf. A1.5.11), and so it is the identity mapping, i.e. the least element of $N(A)$. On the other hand, their join corresponds to the intersection of the corresponding fixsets, i.e. to

$$\{b \in A \mid (a \leq b) \text{ and } (a \Rightarrow b) = b\},$$

but this is the singleton $\{1\}$, i.e. the smallest fixset in A.

(*c*) We call a fixset A_j (or the corresponding nucleus j) *dense* if it has nontrivial intersection with every nontrivial open fixset in A. It is not hard to see that this condition is equivalent to saying that $j(0) = 0$; for if $j(0) = a$ then $A_j \cap A_{o(a)} = \{1\}$, by the argument given above, and conversely if $0 \in A_j$ then $\neg a = (a \Rightarrow 0) \in A_j \cap A_{o(a)}$ for any a. We note that *every frame contains a smallest dense fixset*, namely the subset $A_{\neg\neg} = \{\neg a \mid a \in A\}$: this is a fixset, since it is closed under arbitrary meets ($\neg\neg$ being left adjoint to the inclusion $A_{\neg\neg} \to A$) and an exponential ideal because $(a \Rightarrow \neg b) = \neg(a \wedge b)$ for any a

and b. And $A_{\neg\neg}$ is dense since $\neg\neg 0 = \neg 1 = 0$; but if A_j is any dense fixset then $\neg\neg a = (\neg a \Rightarrow 0) \in A_j$ for all $a \in A$.

(*d*) We shall also occasionally need to invoke the stronger notion of *flat* nucleus: a nucleus j is called flat if it preserves finite joins (including the empty join 0). Equivalently, a fixset A_j is flat if it is a sublattice of A. In general, flat nuclei are much less common than dense ones; we note in particular that the largest dense nucleus $\neg\neg$ is flat iff A is a *Stone algebra*, a condition which we shall discuss in greater detail in Section D4.6.

(*e*) Let (P, T) be a site (as defined in A2.1.9) whose underlying category P is a poset. (Such a site is occasionally referred to as a *posite*.) The subterminal objects in the functor category $[P^{\mathrm{op}}, \mathbf{Set}]$ are those functors F such that each $F(p)$ is either empty or a singleton; so, up to isomorphism, we may identify them with the subsets $\{p \in P \mid F(p) \neq \emptyset\}$, i.e. with the lower subsets of P. The lower sets which correspond to subterminal objects of the subcategory $\mathbf{Sh}(P, T)$ are those sets I with the closure property that, if p is any element of P and $R \in T(p)$ is a cover all of whose members are (more strictly, have their domains) in I, then p itself must be in I. We call such subsets T-*ideals*; the name derives from the fact that, if P is a distributive lattice considered as a coherent category, and T is its coherent coverage, then these subsets are exactly the ideals of P in the classical sense. (Thus the result we are about to establish includes the fact, already mentioned in the proof of 1.1.3, that the ideals of a distributive lattice form a frame.) We write $T\text{-Idl}(P)$ for the set of all T-ideals of P, ordered by inclusion; we claim that it is a fixset in the frame LP of all lower sets in P (so that in particular it is a frame – though we could have deduced that from the fact that we have identified it with the lattice of subterminal objects of a cocomplete topos). It is clearly closed under arbitrary intersections in LP; so we need only verify that it is an exponential ideal. But the Heyting implication in LP is given by

$$(I \Rightarrow J) = \{p \in P \mid (\forall q \leq p)((q \in I) \Rightarrow (q \in J))\},$$

since the latter is the largest lower set whose intersection with I is contained in J. If J is a T-ideal, and $p \in P$ has a covering family R all of whose members are in $(I \Rightarrow J)$, then for any $q \leq p$ such that $q \in I$, we have a covering family R' on q, by the stability condition (C) of A2.1.9, each of whose members lies below some member of R, and is therefore in J since it is in the lower set I. So every such q is in J; hence $p \in (I \Rightarrow J)$, as required.

We note that $T\text{-Idl}(P)$ is a dense fixset in LP (that is, it contains the least element $\emptyset$ of LP) iff every T-covering family is nonempty. (We shall return to this condition in 2.2.4(*e*) below, and again in 3.1.18.) If in addition every T-covering family is (upwards) directed as a subset of P, then $T\text{-Idl}(P)$ is flat in LP; for if I and J are T-ideals and $R \in T(p)$ is a directed cover all of whose members lie in $I \cup J$, then either $R \subseteq I$ or $R \subseteq J$ (if not, we can find $q_1 \in R \setminus I$ and $q_2 \in R \setminus J$, and then an upper bound for $\{q_1, q_2\}$ in R cannot be in either I or J), so $p \in I \cup J$. More generally, if S and T are two coverages on P such that

$S \subseteq T$, and every cover in $T \setminus S$ is directed, then the same argument shows that T-Idl(P) is flat in S-Idl(P). In particular, taking P to be a frame, S to be the coherent coverage on P (i.e. the coverage given by all finite joins in P), and T to be the coverage given by all joins which are either finite or directed, we see that the S-ideals are the ideals in the usual sense and the T-ideals are the principal ideals; thus we have shown that *every frame occurs, up to isomorphism, as a flat fixset in one of the form IA for a distributive lattice A.*

We remark that the use of non-constructive logic in the preceding paragraph cannot be avoided: the complementations can be eliminated, but if this is done the argument is then seen to depend on an instance of the 'dual Frobenius rule' that $(\phi \vee (\forall x)\psi(x))$ may be deduced from $(\forall x)(\phi \vee \psi(x))$ – cf. D1.3.9. However, the final italicized result is constructive: it is easy to see directly that any frame A is both a sublattice of, and a fixset in, its ideal lattice IA.

A further important tool in elucidating the structure of $N(A)$ is the following:

Lemma 1.1.17 *Any nucleus on a frame is expressible as a join of nuclei of the form $o(a) \wedge c(b)$.*

Proof Let A be a frame, and j an arbitrary nucleus on A. For each $a \in A$, consider the nucleus $k_a = o(a) \wedge c(j(a))$. We have

$$k_a(a) = (a \Rightarrow a) \wedge (j(a) \vee a) = 1 \wedge j(a) = j(a) .$$

But for any b, we have

$$(a \Rightarrow b) \wedge j(a) \leq j(a \Rightarrow b) \wedge j(a) = j((a \Rightarrow b) \wedge a) \leq j(b),$$

so that

$$\begin{aligned} k_a(b) &= (a \Rightarrow b) \wedge (j(a) \vee b) \\ &= ((a \Rightarrow b) \wedge j(a)) \vee ((a \Rightarrow b) \wedge b) \\ &\leq j(b) \vee b = j(b) . \end{aligned}$$

Thus j is the (pointwise) join of the nuclei k_a, $a \in A$. □

Next, we investigate the functoriality of the construction $A \mapsto N(A)$. For this we need

Lemma 1.1.18 *Let $h\colon A \to B$ be a frame homomorphism, $a \in A$. Then*

(i) *h factors through $A \to A_{o(a)}$ iff $h(a) = 1_B$.*
(ii) *h factors through $A \to A_{c(a)}$ iff $h(a) = 0_B$.*

(iii) *There are pushout squares*

$$\begin{array}{ccc} A & \xrightarrow{h} & B \\ \downarrow & & \downarrow \\ A_{o(a)} & \longrightarrow & B_{o(h(a))} \end{array} \qquad \textit{and} \qquad \begin{array}{ccc} A & \xrightarrow{h} & B \\ \downarrow & & \downarrow \\ A_{c(a)} & \longrightarrow & B_{c(h(a))} \end{array}$$

in **Frm**.

Proof (i) One direction is immediate, since a maps to the top element of $A_{o(a)}$. Conversely, suppose $h(a) = 1_B$. Then for any $b \in A$, we have

$$h(a \Rightarrow b) \leq h(a) \Rightarrow h(b) = 1 \Rightarrow h(b) = h(b)$$

(where the first inequality derives from the fact that h preserves finite meets); so h identifies any two elements of A which are identified by $o(a)$.

(ii) is similar, using the fact that a maps to the bottom element of $A_{c(a)}$.

(iii) follows immediately from (i) and (ii). □

Proposition 1.1.19 *The assignment $A \mapsto N(A)$ defines a functor* **Frm** $\to$ **Frm**; *and the frame homomorphisms $c\colon A \to N(A)$ define a natural transformation from the identity to this functor, which is both monic and epic.*

Proof Given a frame homomorphism $h\colon A \to B$, the operation of pushing out maps along h preserves regular epimorphisms, and so defines a mapping $N(A) \to N(B)$, which we take to be $N(h)$. It requires verification that $N(h)$ is a frame homomorphism; but it clearly preserves joins, since they correspond to cointersections of regular quotients, and so we need only consider finite meets.

First we consider a meet of the form $o(a) \wedge c(b)$. From 1.1.18(iii), we have $N(h)(o(a)) = o(h(a))$ and $N(h)(c(b)) = c(h(b))$; and since $o(a) \wedge c(b)$ is complementary to $c(a) \vee o(b)$, by 1.1.16(*b*), a straightforward calculation shows that $N(h)(o(a) \wedge c(b)) = o(h(a)) \wedge c(h(b))$. Also, since $o(a_1) \wedge o(a_2) = o(a_1 \vee a_2)$ and $c(b_1) \wedge c(b_2) = c(b_1 \wedge b_2)$, the meet of two nuclei of this form is again of the same form, and $N(h)$ preserves all such meets. The general case now follows from 1.1.17: by this result, plus the infinite distributive law in $N(A)$, we can reduce an arbitrary meet $j \wedge k$ to a join of pairwise meets of nuclei of the form $o(a) \wedge c(b)$, and hence it is preserved by $N(h)$.

Next, we observe that c is a frame homomorphism: it preserves finite meets by distributivity of A, and arbitrary joins because joins commute with joins – note that joins of closed nuclei are actually computed pointwise. The fact that c is a natural transformation is immediate from 1.1.18(iii); it is monic because we can recover a from $c(a)$ by applying it to the least element of A, and epic because we saw in 1.1.17 that $N(A)$ is generated as a frame by the elements $c(a)$ and their complements (and frame homomorphisms preserve complements). □

Corollary 1.1.20

(i) *For any frame A, $c\colon A \to N(A)$ is universal among frame homomorphisms $h\colon A \to B$ mapping all elements of A to complemented elements of B.*

(ii) *$c\colon A \to N(A)$ is an isomorphism iff A is a complete Boolean algebra.*

Proof (i) By 1.1.16(b), we know that c has this property. Suppose given another such homomorphism h. Then, for any $a \in A$, $N(h)(c(a)) = c(h(a))$ is a closed nucleus on B; but so is $N(h)(o(a))$, since $o(b_1) = \neg c(b_1) = c(b_2)$ if b_1 and b_2 are complementary elements of B. It now follows from 1.1.17 that $N(h)$ maps arbitrary nuclei on A to closed nuclei on B, i.e. it factors through $c\colon B \to N(B)$. This factorization is a frame homomorphism $N(A) \to B$ extending h; and it is unique, since $c\colon A \to N(A)$ is epic.

(ii) It is clear that the identity morphism on A has the universal property mentioned in (i) iff all elements of A are already complemented, i.e. iff A is a Boolean algebra. □

Remark 1.1.21 Using 1.1.20, we might try to construct a left adjoint for the inclusion functor **CBool** $\to$ **Frm**, where **CBool** is the category of complete Boolean algebras, as follows. Given a frame A, we may define a sequence of frames $N_\alpha(A)$ for all ordinals α, and frame homomorphisms $c_{\alpha,\beta}\colon N_\alpha(A) \to N_\beta(A)$ for all $\alpha \leq \beta$, by recursion:

(i) $N_0(A) = A$, $c_{0,0} = 1_A$ (in fact $c_{\alpha,\alpha}$ is the identity for all α);

(ii) $N_{\alpha+1}(A) = N(N_\alpha(A))$, $c_{\beta,\alpha+1} = c \cdot c_{\beta,\alpha}$ (where the unsubscripted 'c' is the homomorphism of 1.1.19);

(iii) if λ is a nonzero limit ordinal, then $N_\lambda(A)$ is the colimit in **Frm** of the diagram formed by the $N_\alpha(A)$ and the $c_{\alpha,\beta}$ for $\alpha, \beta < \lambda$, and the $c_{\alpha,\lambda}$ are the legs of the colimiting cone.

An easy extension of 1.1.20 shows that any frame homomorphism $h\colon A \to B$, where B is a complete Boolean algebra, factors uniquely through $c_{0,\alpha}\colon A \to N_\alpha(A)$ for all α; hence *if there exists α such that $N_\alpha(A)$ is Boolean (equivalently, such that $c_{\alpha,\alpha+1}$ is an isomorphism), then $N_\alpha(A)$ is the reflection of A in* **CBool**.

However, in general this reflection does not exist: for it is well known that the forgetful functor **CBool** $\to$ **Set** has no left adjoint (specifically, the free complete Boolean algebra on a countable infinity of generators does not exist), and so if we take A to be the free frame on a countable infinity of generators, i.e. the open-set lattice of the countable power of Sierpiński space (cf. 1.1.4), then the frames $N_\alpha(A)$ are all non-Boolean. But it is also easy to see that the morphisms $c_{0,\alpha}$ are epic (as well as monic) for all α: given $h, k\colon N_\alpha(A) \rightrightarrows B$ such that $h \cdot c_{0,\alpha} = k \cdot c_{0,\alpha}$, we may prove by induction on β that $h \cdot c_{\beta,\alpha} = k \cdot c_{\beta,\alpha}$ for all $\beta \leq \alpha$. So we conclude that *the frame $A = \mathcal{O}(S^{\mathbb{N}})$ has a proper class of epimorphic images in* **Frm**, *no two of which are isomorphic in the co-slice*

category $A\backslash\mathbf{Frm}$. In particular, the category **Frm** is not well-copowered, and it does not have a cogenerating set of objects (cf. A1.4.17).

Finally in this section, we briefly introduce the 'fibrewise' versions of denseness and closedness for nuclei, which we shall need for our development of the theory of localic groups in Chapter C5. Given a frame homomorphism $h: A \to B$, we say that a nucleus j on B is *fibrewise dense* relative to A if $j(h(a)) = h(a)$ for all $a \in A$. We also say j is *fibrewise closed* relative to A if is the least nucleus in its A-equivalence class, where two nuclei are called A-equivalent if they agree on all elements of the form $h(a)$, $a \in A$. The reason for the names is that in the 'classical' case when h is the frame homomorphism $\mathcal{O}(X) \to \mathcal{O}(Y)$ induced by a continuous map of spaces $f: Y \to X$, and the nucleus j corresponds to a subspace Y' of Y, then j is fibrewise dense (resp. fibrewise closed) relative to $\mathcal{O}(X)$ iff $Y' \cap f^{-1}(x)$ is dense (resp. closed) in the fibre $f^{-1}(x)$, for each point $x \in X$. (We shall prove this in 1.2.16 below.) However, we do not have a corresponding notion of 'fibrewise open nucleus', as might have been expected; fibrewise closed nuclei are not in general complemented in $N(B)$. (We may see this by taking h to be the identity map $B \to B$, in which case every nucleus on B becomes fibrewise closed.)

Since f preserves the bottom element, it is clear that fibrewise dense implies dense in the sense of 1.1.16(c), and similarly closed in the sense of 1.1.16(b) implies fibrewise closed, since $c(b)$ is the least nucleus sending 0 to b. If A is the two-element frame $\{0, 1\}$, then the converse implications hold, since any nucleus fixes 1; however, constructively it is not true that the converse implications hold even for fibrewise closedness and denseness relative to the initial frame Ω – we shall sometimes use the terms 'weakly closed' and 'strongly dense' for the fibrewise notions relative to Ω, to distinguish them from the notions defined in 1.1.16. We note the following elementary properties of fibrewise closedness and denseness:

Lemma 1.1.22 *Let $h: A \to B$ be a frame homomorphism. Then*

(i) *There exists a largest A-fibrewise dense nucleus on B.*

(ii) *For any nucleus j on B, there exists an A-fibrewise closed nucleus j^A (its* A-fibrewise clósure*) which agrees with j on elements of the form $h(a)$, $a \in A$.*

(iii) *The A-fibrewise closed nuclei on B form a subframe $N_A(B)$ of $N(B)$, and there is a pushout square*

$$\begin{tikzcd} A \arrow[r, "c"] \arrow[d, "h"] & N(A) \arrow[d] \\ B \arrow[r] & N_A(B) \end{tikzcd}$$

in **Frm**.

Proof Since the notions depend only on the image of $h: A \to B$, we shall simplify our notation by assuming that h is injective and identifying A with a subframe of B. (Indeed, for much of this proof, we could generalize to the case when A is an arbitrary subset of B; but we shall not need this extra generality.)

(i) We have to show that an arbitrary join of fibrewise dense nuclei is fibrewise dense. But joins of nuclei correspond to intersections of fixsets, as we observed in the proof of 1.1.15; and it is clear that j is fibrewise dense iff the fixset B_j contains A. So the result is immediate.

(ii) Clearly, we may define j^A to be the meet of all nuclei agreeing with j on the subframe A, since meets of nuclei are constructed pointwise. Alternatively, we may define j^A to be the join

$$\bigvee\{(o(a) \wedge c(j(a))) \mid a \in A\},$$

and verify as in the proof of 1.1.17 that we have $j^A(b) \leq j(b)$ for all b, with equality if $b \in A$. But this definition of j^A depends only on the A-equivalence class of j, as defined above; so j^A is the least element of this equivalence class.

(iii) We claim that $N_A(B)$ is exactly the subframe of $N(B)$ generated by all elements of the form $o(a)$ $(a \in A)$ or $c(b)$ $(b \in B)$. It is clear from the second construction of the fibrewise closure in (ii) that every fibrewise closed nucleus belongs to this subframe; for the reverse inclusion, it suffices to show that any nucleus of the form $o(a) \wedge c(b)$ is fibrewise closed, since the set of nuclei of this form is closed under finite meets, and it is clear that joins of fibrewise closed nuclei are fibrewise closed (since, by (ii), they form a coreflective sub-poset of $N(B)$). But if j is any nucleus agreeing with $o(a) \wedge c(b)$ on elements of A, then in particular $j(a) = (a \vee b)$, whence for any b' we have

$$j(b') \geq j(a) \wedge j(a \Rightarrow b') \geq b \wedge (a \Rightarrow b');$$

and $j(b') \geq b'$ in any case, so $j(b') \geq (b \vee b') \wedge (a \Rightarrow b') = (c(b) \wedge o(a))(b')$.

In particular, we have shown that $N_A(B)$ is a subframe of $N(B)$. To show that the diagram is a pushout, it suffices to note that this subframe contains the image of $c: B \to N(B)$, and that a frame homomorphism $k: B \to C$ factors (uniquely) through $B \to N_A(B)$ iff $k(a)$ (that is, $k(h(a))$) is complemented in C for all $a \in A$, by the argument in the proof of 1.1.20(i). □

We note that the fibrewise closure j^A of j may be characterized as the unique fibrewise closed nucleus k on B such that B_j is fibrewise dense as a fixset in B_k. We shall derive some further elementary properties of these notions in 1.2.14 below.

Suggestions for further reading: Johnstone [520], Johnstone & Vickers [548], Simmons [1110, 1111], Vajner [1182].

C1.2 Locales and spaces

In the previous section, we studied frames from an 'algebraic' point of view. However, it is important to recognize that they have topological aspects as well: we are primarily interested in thinking of frames *as if* they were lattices of the form $\mathcal{O}(X)$ for a topological space X, even though in fact they are not all of this form. For this reason, it is convenient to introduce a different terminology and notation, which we shall use when thinking of frames from a topological point of view.

Definition 1.2.1 A *locale* is the same thing as a frame; however, we use letters like $X, Y, \ldots$ to denote locales, and write $\mathcal{O}(X)$ for the frame corresponding to a locale X. Elements of $\mathcal{O}(X)$ will be denoted by letters like $U, V, \ldots$. A *continuous map* of locales $f: X \to Y$ is defined to be a frame homomorphism $f^*: \mathcal{O}(Y) \to \mathcal{O}(X)$. We write **Loc** for the category of locales – that is, for the category $\mathbf{Frm}^{\mathrm{op}}$.

Thus the functor $\mathcal{O}: \mathbf{Sp} \to \mathbf{Frm}^{\mathrm{op}}$ becomes nameless as a functor $\mathbf{Sp} \to \mathbf{Loc}$: if X is a space, we write X also for the locale which 'is' the frame $\mathcal{O}(X)$ of open subsets of X. (If we allow consideration of non-sober spaces – which we shall shortly legislate out of existence – this notation is potentially ambiguous; for the moment, we shall remove the ambiguity by writing X_l for the locale corresponding to a space X.)

We define a *point* of a locale X to be a continuous map $p: 1 \to X$, where 1 denotes (the locale corresponding to) the one-point space; equivalently, it is a frame homomorphism $p^*: \mathcal{O}(X) \to \mathcal{O}(1)$. Such a map p^* is determined by the subset $F = \{U \in \mathcal{O}(X) \mid p^*(U) = 1\}$ of $\mathcal{O}(X)$; the requirement that p^* be a frame homomorphism corresponds to the requirement that F be a *completely prime filter* in $\mathcal{O}(X)$, that is a filter such that $\bigvee S \in F$ implies $S \cap F \neq \emptyset$, for any $S \subseteq \mathcal{O}(X)$. So we frequently identify points of X with such filters; that is, instead of regarding open sets as sets of points (as is usual in topology), locale theory regards points as (appropriate) sets of 'open sets'. (There is clearly scope for a more symmetric development of topology, in which we take both 'point' and 'open set' as primitive notions, together with a binary relation between them. For an indication of this line of development, see [1205].) We note that if X is a space and p a point of X (in the classical sense), then the completely prime filter in $\mathcal{O}(X)$ which corresponds to p is simply $\{U \in \mathcal{O}(X) \mid p \in U\}$, the *neighbourhood filter* of p.

Classically, a filter is completely prime iff its complement is a principal prime ideal; so it is tempting to identify the points of a locale with the prime elements of the corresponding frame (i.e. those which generate prime principal ideals). However, we shall do our best to resist this temptation, since we wish as many as possible of our arguments concerning locales to be constructively valid, and so applicable to internal locales in an arbitrary topos. (There are some results – for

example 1.2.4 below – where we shall have to employ non-constructive reasoning; we shall try always to indicate these explicitly.)

Writing X_p for the set of all points of a locale X, we have a map $\phi_X : \mathcal{O}(X) \to P(X_p)$ (where $P(A)$ denotes the power-set of A), given by $\phi_X(U) = \{p \in X_p \mid p^*(U) = 1\}$. It is easy to verify that ϕ_X is a frame homomorphism, and so its image is a topology on X_p; from now on, we regard X_p as a space equipped with this topology. Moreover, if $f : X \to Y$ is a continuous map of locales, then composition with f^* defines a mapping $X_p \to Y_p$, which is easily seen to be continuous with respect to the topologies just defined; that is, $X \mapsto X_p$ becomes a functor **Loc** $\to$ **Sp**.

Lemma 1.2.2 *The functor $X \mapsto X_p$ just defined is right adjoint to the functor $X \mapsto X_l$:* **Sp** $\to$ **Loc***. Moreover, the adjunction is idempotent.*

Proof For a space X, the mapping η_X which sends a point p of X to the completely prime filter $\{U \in \mathcal{O}(X) \mid p \in U\}$ is a continuous map $X \to X_{lp}$, since we have $\eta_X^{-1}(\phi_{X_l}(U)) = U$ for any $U \in \mathcal{O}(X)$. And for a locale Y, we have a continuous map $\epsilon_Y : Y_{pl} \to Y$ given by $\epsilon_Y^* = \phi_Y : \mathcal{O}(Y) \to \mathcal{O}(Y_p)$. These maps are easily verified to be natural in X and Y, and to satisfy the triangular identities for an adjunction. We note that ϕ_Y is always surjective, by the definition of the topology on Y_p; if Y is a locale of the form X_l, then it is bijective, since if U and V are distinct open subsets of X then there must be a point of X lying in just one of them, and composition with η_X yields a point of Y lying in just one of $\phi_Y(U)$ and $\phi_Y(V)$. Thus ϵ_{X_l} is an isomorphism for all spaces X, i.e. the adjunction is idempotent. (It follows from the general theory of adjunctions that η_{Y_p} must be an isomorphism for all locales Y; in fact this could have been proved directly, by a similar argument.) □

We call a space X *sober* if η_X is an isomorphism; in fact, given the way in which the topology is defined on X_{lp}, this is equivalent to demanding that η_X be bijective, i.e. that each completely prime filter in $\mathcal{O}(X)$ is of the form $\{U \in \mathcal{O}(X) \mid p \in U\}$ for a *unique* $p \in X$. And we call a locale Y *spatial* if ϵ_Y is an isomorphism. We thus obtain three immediate corollaries of 1.2.2:

Corollary 1.2.3

(i) *The full subcategory* **Sob** *of* **Sp** *whose objects are sober spaces is reflective in* **Sp***, the reflector being given by $X \mapsto X_{lp}$. We call X_{lp} the* sobrification *of X.*

(ii) *The full subcategory* **SLoc** *of* **Loc** *consisting of spatial locales is coreflective, the coreflector being given by $Y \mapsto Y_{pl}$.*

(iii) *The adjunction of* 1.2.2 *restricts to an equivalence of categories between* **Sob** *and* **SLoc***.* □

In passing, we note that the T_0 axiom for spaces is equivalent to the assertion that η_X is injective, since $\eta_X(p) = \eta_X(q)$ iff there is no open set in Xcontaining

just one of p and q. Thus sober spaces are necessarily T_0. In the opposite direction, we have

Lemma 1.2.4 *Hausdorff spaces are sober.*

Proof Let X be a Hausdorff space, and F a completely prime filter in $\mathcal{O}(X)$. Then

$$V = \bigcup\{U \mid U \in (\mathcal{O}(X) \setminus F)\}$$

is an open set which is not the whole of X (since it is not in F) but has the property that any open set strictly containing it must be in F. Now if x and y are two distinct points of $X \setminus V$, then we can find disjoint open neighbourhoods M, N of x and y; and then $V \cup M$ and $V \cup N$ are open sets in F whose intersection is not in F, contradicting the fact that F is a filter. So $X \setminus V$ is a singleton $\{x\}$, from which it follows easily that $F = \{U \mid x \in U\}$. □

We remark that the use of complementation in the above proof is unavoidable; unlike the other results proved so far in this chapter, 1.2.4 is not constructively valid. We shall give a counterexample in 1.6.8 below.

There is no implication in either direction between sobriety and the T_1 axiom. If X is an infinite set equipped with the cofinite topology, then the family of all nonempty open subsets of X is a completely prime filter which is not in the image of η_X; so X is not sober. And the Sierpiński space S is easily seen to be sober but not T_1. (It is also of interest to note that the space KA (with topology $L(KA^{\mathrm{op}})$) considered in the proof of 1.1.4 is not sober if A is infinite: the completely prime filters of $L(KA^{\mathrm{op}})$ correspond bijectively to arbitrary subsets of A, not just to finite ones – and in fact the sobrification of KA turns out to be homeomorphic to S^A.)

Despite the foregoing examples, it seems to be the case that one does not lose any interesting examples of topological spaces by restricting one's attention to sober spaces. (Essentially, this is because most of the interesting topological properties of a space – such as compactness and connectedness – reside in its open-set lattice; and we can always replace a non-sober space by a sober one having the same open-set lattice, namely its sobrification.) So from now on, unless we specifically say otherwise, all spaces will be assumed to be sober; and we shall drop the use of the notation X_l – a (sober) space will be identified with the spatial locale to which it corresponds. (Thus the spatial coreflection of a locale Y will simply be denoted by Y_p, not Y_{pl}; and if we ever do encounter a non-sober space we shall denote its sobrification by X_s rather than X_{lp}. The point is that it really doesn't matter, for a lot of the things we wish to do, whether we consider a particular object to be a sober space or a spatial locale.)

In passing from spaces to locales, we have thus sacrificed the 'pathological' non-sober spaces, and gained the non-spatial locales. The reader may be forgiven for wondering whether this is a genuine gain: are not the non-spatial locales (of which, be it noted, we haven't given any examples yet) just as pathological as the

non-sober spaces? Before we can give a straight answer to this question (cf. 1.2.8 and 1.2.9 below), we need to consider sublocales, i.e. regular monomorphisms in **Loc**.

We define a *sublocale* of a locale X to be a regular subobject of X in **Loc**; i.e. it is the locale corresponding to a regular quotient of $\mathcal{O}(X)$. We write $\mathrm{Sub}(X)$ for the lattice of sublocales of X; note that the natural ordering on it (defined as in Section A1.3) corresponds to that on fixsets in $\mathcal{O}(X)$, so as a lattice $\mathrm{Sub}(X)$ is isomorphic to $N(\mathcal{O}(X))^{\mathrm{op}}$. (This notation conflicts with our usual practice, introduced in Section A1.3, whereby $\mathrm{Sub}(X)$ denotes the preorder of *all* subobjects of X; but we shall not have occasion to consider this latter concept in the category of locales – one reason is that **Loc** is not well-powered, as we saw in 1.1.21. We shall also use the phrase 'image factorization' in **Loc** in the opposite of our usual sense, introduced in Section A1.3: that is, the image factorization of a continuous map of locales is its factorization into an epimorphism followed by a regular monomorphism, corresponding to the surjection–injection factorization of the corresponding frame homomorphism. Again, this nonstandard usage may be justified by the fact that **Loc** is not regular; in fact its regular epimorphisms are not even stable under composition – see [980].)

We note also that if Y is a subspace of a space X, then the definition of subspace topology implies that $i^{-1}\colon \mathcal{O}(X) \to \mathcal{O}(Y)$ is a surjective frame homomorphism (where $\imath\colon Y \to X$ is the inclusion); so the sublocales of a space include all its subspaces. Incidentally, we should mention that not all subspaces of a sober space need be sober; but we have

Lemma 1.2.5 *Let X be a (not necessarily sober) space. Then there is a closure operation $Y \mapsto \check{Y}$ on subsets of X, such that for any Y, $\check{Y}$ is the largest subspace for which the inclusion $Y \rightarrowtail \check{Y}$ induces a bijection of open-set lattices. If X is sober, then $\check{Y}$ is (homeomorphic to) the sobrification of Y, for any $Y \subseteq X$.*

Proof We recall that a subset of X is said to be *locally closed* if it is expressible as the intersection of an open and a closed subset. We define $\check{Y}$ (which is sometimes called the *subclosure* of Y) to be the set of all points x such that every locally closed set containing x meets Y.

Now if U and V are open sets in X having the same intersection with Y, then they have the same intersection with $\check{Y}$; for the locally closed sets $U \setminus V$ and $V \setminus U$ do not meet Y, and hence cannot meet $\check{Y}$. Thus $\mathcal{O}(Y) \cong \mathcal{O}(\check{Y})$. On the other hand, if x is a point of $X \setminus \check{Y}$, then it belongs to the difference of two open sets having the same intersection with Y, and so cannot belong to any subspace having the same open-set lattice as Y. So $\check{Y}$ is the largest subspace of X with this property.

Now suppose X is sober, and let F be a completely prime filter of open sets in Y. Then $G = \{U \in \mathcal{O}(X) \mid U \cap Y \in F\}$ is a completely prime filter in $\mathcal{O}(X)$, and so defines a unique point x of X such that $U \in G$ iff $x \in U$. Now if $C = U \cap (X \setminus V)$ is a locally closed set containing x, then $U \in G$ but $V \notin G$, so $U \cap Y \not\subseteq V \cap Y$, i.e. $C \cap Y \neq \emptyset$. So $x \in \check{Y}$. Conversely, if $x \in \check{Y}$,

then $\{U \in \mathcal{O}(\check{Y}) \mid x \in U\}$ is a completely prime filter in $\mathcal{O}(\check{Y})$, and so the intersections of these sets with Y form a completely prime filter in $\mathcal{O}(Y)$. Thus we have set up a bijection between the points of $\check{Y}$ and those of Y_s; it is easy to see that it must be a homeomorphism. □

A space X is said to be a *T_D-space* if every singleton subset of X is locally closed. For such spaces, it is clear that every subspace Y is subclosed (i.e. satisfies $Y = \check{Y}$), and so *every subspace of a sober T_D-space is sober.* Note also that open subspaces and closed subspaces of sober spaces are always sober.

Reverting to our main theme, we now know that if X is a (sober) space then $\mathrm{Sub}(X)$ contains a copy of the poset of subclosed subsets of X. (However, they do not form a sublattice of $\mathrm{Sub}(X)$: unions of spatial sublocales are spatial, but we shall see that intersections need not be.) Next, we revisit the first three examples of nuclei studied in 1.1.16, to indicate our localic notation for them.

Examples 1.2.6 (*a*) Let X be a locale, $U \in \mathcal{O}(X)$. Then we have a sublocale of X corresponding to the open nucleus $o(U)$. We denote this sublocale by U itself; in other words, we identify $\mathcal{O}(X)$ with the poset of open sublocales of X. This identification is 'harmless' because the mapping $o\colon \mathcal{O}(X) \to N(\mathcal{O}(X)) \cong \mathrm{Sub}(X)^{\mathrm{op}}$ sends finite meets to joins and arbitrary joins to meets, using 1.1.19 and the fact that open nuclei are complementary to closed ones. In keeping with it, we shall henceforth tend to denote the top and bottom elements of $\mathcal{O}(X)$ by X and $\emptyset$ respectively, rather than 1 and 0; and we shall tend to use $\bigcup$ and $\cap$, rather than $\bigvee$ and $\wedge$, to denote joins and finite meets in $\mathcal{O}(X)$. Note that if X is a space, and U is an open subset of X, we may identify the sublocale U with the space obtained by equipping U with the subspace topology, since the open-set lattice of the latter may be identified with the principal ideal $\downarrow(U)$ in $\mathcal{O}(X)$. In particular, *an open sublocale of a spatial locale is spatial.*

(*b*) For the closed sublocale of X corresponding to the nucleus $c(U)$ on $\mathcal{O}(X)$, we shall use the notation $\complement U$. Once again, if X is a space and U an open subset of X, we may identify this sublocale with the space obtained by equipping the closed set $X \setminus U$ with the subspace topology; so *a closed sublocale of a spatial locale is spatial.* (However, the use of complementation is an indication that this result, unlike that cited at the end of (*a*), is not constructively valid. Indeed, the assertion 'every closed sublocale of a discrete locale is spatial' implies the law of excluded middle: for in a discrete locale (that is, one corresponding to a frame of the form PX) every point is an open sublocale, and hence every spatial sublocale (being the union of its points) is open. The quoted assertion thus implies that the lattice of open sublocales of any discrete locale (that is, PX for any X) is Boolean.)

(*c*) Of course, we call a sublocale of X *dense* if it corresponds to a dense nucleus on $\mathcal{O}(X)$. It is again easy to see that any dense subspace of a space X gives rise to a dense sublocale of X; but in general a space will have many more dense sublocales than dense subspaces. In particular, we write X_b for the smallest dense sublocale of X, corresponding to the nucleus $\neg\neg$ ('b' stands for

'Boolean': recall that $\mathcal{O}(X_b) = (\mathcal{O}(X))_{\neg\neg}$ is a Boolean algebra). It is rarely spatial: for example, if X is a sober T_D-space without isolated points (i.e. such that $\{x\}$ is not open for any $x \in X$), then the subspaces $X \setminus \{x\}$ are all dense, so X_b is contained in their intersection, and thus cannot have any points.

The fact that intersections of dense sublocales are dense is thus the key difference between the categories **Sob** and **Loc**. In fact, **Loc** is in a sense the minimal modification of **Sob** which makes this result possible:

Lemma 1.2.7 *Assuming classical logic, any locale is expressible as an intersection of dense spatial sublocales of a spatial locale.*

Proof We saw in 1.1.4 that free frames are spatial locales; and every frame is a quotient of the free frame on its underlying set, i.e. any locale is a sublocale of a spatial locale. To obtain the stronger statement above, note first that any sublocale has a closure: that is, a regular monomorphism $Y \rightarrowtail X$ in **Loc** has a factorization $Y \rightarrowtail \overline{Y} \rightarrowtail X$, where $Y \rightarrowtail \overline{Y}$ is dense and $\overline{Y} \rightarrowtail X$ is closed. (If j is the nucleus on $\mathcal{O}(X)$ corresponding to Y, then $\overline{Y}$ corresponds to $c(j(\emptyset))$.) Since closed sublocales of spatial locales are spatial, we now know that any locale Y is expressible as a dense sublocale of a spatial locale X. Lemma 1.1.17 then expresses it as an intersection of sublocales of the form $U \cup \complement V$, which are spatial because they are unions of spatial sublocales of X, and dense in X because they contain Y. $\square$

We now give a couple of interesting examples of non-spatial locales, which we hope may convince even the classically-minded reader that something has been gained by the switch from spaces to locales.

Example 1.2.8 Let A be a nonempty set, and let X be the set of all functions $\mathbb{N} \to A$, equipped with the Tychonoff topology where A is regarded as a discrete space, i.e. the topology in which the basic open neighbourhoods of a function f have the form

$$\{g \in X \mid g|_n = f|_n\}$$

for some $n \geq 0$ (where $g|_n$ means the restriction of g to $\{0, 1, \ldots, n-1\} \subseteq \mathbb{N}$). Note that X is (a Hausdorff space, and therefore) sober. For each $a \in A$, let X_a be the subspace $\{f \in X \mid a \in \text{im } f\}$; clearly X_a is dense in X. Let $Y = \bigcap_{a \in A} X_a$, where the intersection is computed in the lattice of sublocales of X; then Y is also dense in X. However, Y_p is the intersection of the X_a in the lattice of subspaces of X; so if A is uncountable then $Y_p = \emptyset$, and hence Y is not spatial.

We may identify $\mathcal{O}(X)$ with the Lindenbaum algebra (cf. D1.4.14) of the geometric propositional theory of functions $\mathbb{N} \to A$, that is the theory over the signature with primitive propositions $p(n,a)$ indexed by $\mathbb{N} \times A$ whose axioms are

$$\begin{array}{ll} (\top \vdash \bigvee_{a \in A} p(n,a)) & \text{for all } n, \text{ and} \\ ((p(n,a) \wedge p(n,b)) \vdash \bot) & \text{for all } n, a, b \text{ with } a \neq b. \end{array}$$

If we add the further axioms

$$(\top \vdash \bigvee_{n \in \mathbb{N}} p(n,a)) \quad \text{for all } a,$$

we obtain the geometric theory of surjections $\mathbb{N} \twoheadrightarrow A$, whose Lindenbaum algebra is $\mathcal{O}(Y)$. The fact that this theory is always consistent, even though it may have no models in **Set** (i.e. Y may have no points) will be of importance in the representation theorems of Chapter C5.

Example 1.2.9 For another example in a similar vein, consider the localic translation of 1.1.12: *if $(X_i \mid i \in I)$ are the vertices of a cofiltered diagram in* **Loc** *whose transition maps $X_j \to X_i$ are all epimorphisms, then the legs $(\varprojlim_{i \in I} X_i) \to X_i$ of the limiting cone are also epic.* This is well known to be false for spaces (even for discrete spaces): again, let A be an uncountable set, let I be the poset of finite subsets of A ordered by reverse inclusion, for each $i \in I$ let X_i be the set of injections $i \rightarrowtail \mathbb{N}$ (with the discrete topology), and for $j \supseteq i$ let $X_j \to X_i$ be the restriction map. Then a point of the inverse limit $\varprojlim_{i \in I} X_i$ in **Sp** would be an injection $A \rightarrowtail \mathbb{N}$, so no such points exist; but the inverse limit in **Loc** is nontrivial, since it maps epimorphically to each X_i.

Once again, we may think of the frame corresponding to this locale as a Lindenbaum algebra: this time for the theory of injections $A \rightarrowtail \mathbb{N}$, which we may present using the same signature as the theory of 1.2.8, with the second and third groups of axioms displayed there, plus

$$((p(m,a) \wedge p(n,a)) \vdash \bot) \quad \text{for all } (m,n,a) \text{ with } m \neq n.$$

The reader may be prompted to ask at this point whether the theory of bijections $A \to \mathbb{N}$ (i.e. the propositional theory with all four groups of axioms from 1.2.8 and 1.2.9) also has a nontrivial Lindenbaum algebra when A is uncountable. The answer is yes, but it seems harder to give a direct locale-theoretic proof of this with the tools currently at our disposal. (One possible proof runs as follows: since the theory of injections $A \rightarrowtail \mathbb{N}$ is consistent by 1.2.9, it has a model in a non-degenerate Boolean topos $\mathcal{B}$ by D3.1.16; but in a Boolean topos the Cantor–Bernstein theorem is valid by D4.1.11, and so we can construct a bijection $A \to \mathbb{N}$ in $\mathcal{B}$.) We shall also encounter the intersection of the theories of 1.2.8 and 1.2.9 in 3.3.11(c) below (and see also D4.1.9).

Restating Lemma 1.1.18(iii) in localic terms, we have

Lemma 1.2.10 *Let $f\colon X \to Y$ be a continuous map of locales, $U \in \mathcal{O}(Y)$. Then there are pullback squares*

$$\begin{tikzcd} f^*(U) \arrow[r] \arrow[d, tail] & U \arrow[d, tail] \\ X \arrow[r, "f"] & Y \end{tikzcd} \quad \textit{and} \quad \begin{tikzcd} \mathsf{C}(f^*(U)) \arrow[r] \arrow[d, tail] & \mathsf{C}U \arrow[d, tail] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

in **Loc**. □

The 'open' half of 1.2.10 tells us that there is no ambiguity, given our identification of elements of $\mathcal{O}(X)$ with open sublocales of X, in using f^* both for the frame homomorphism corresponding to a continuous map $f\colon X \to Y$ of locales and (as usual) for the effect of pulling back sublocales of Y along f. We shall write $f_!\colon \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ for the left adjoint of f^*, i.e. the map which sends a sublocale S of X to the image of the composite $S \rightarrowtail X \to Y$. (In general, $f_!$ does not restrict to a map $\mathcal{O}(X) \to \mathcal{O}(Y)$; we shall investigate the circumstances in which it does so in Section C3.1.)

Although sublocales are the 'natural' localic generalization of subspaces, the fact that they fail to form a Boolean algebra is sometimes inconvenient. For some purposes, it is useful to restrict our attention to the Boolean algebra $\mathrm{Sub}(X)^c$ of complemented sublocales of X: note that this contains all open or closed sublocales of X. (However, it is rarely complete as a lattice.) For future reference, we note a couple of results concerning complemented sublocales.

Lemma 1.2.11 *Let $f\colon X \to Y$ be an epimorphism in* **Loc**. *Then every complemented sublocale S of Y satisfies $S = f_! f^*(S)$. In particular, f^* is injective on complemented sublocales.*

Proof Let T be the complement of S in $\mathrm{Sub}(Y)$. The maps f^* and $f_!$ both preserve finite joins: $f_!$ because it is a left adjoint, and f^* because it is a coframe homomorphism (by 1.1.19). So $f_! f^*(S) \cup f_! f^*(T) = f_! f^*(Y) = f_!(X) = Y$ (the last step using the fact that f is epic). But from the adjunction we have $f_! f^*(S) \leq S$ and $f_! f^*(T) \leq T$, so

$$\begin{aligned} S &= S \cap (f_! f^*(S) \cup f_! f^*(T)) \\ &= (S \cap f_! f^*(S)) \cup (S \cap f_! f^*(T)) \\ &\leq f_! f^*(S) \cup (S \cap T) \\ &= f_! f^*(S) \cup \emptyset \,. \end{aligned}$$

The second assertion is immediate from the first. □

Example 1.2.12 To show that the result of 1.2.11 fails for arbitrary sublocales, let Y be a sober T_D-space without isolated points, let X be the same set retopologized with the discrete topology, and let $f\colon X \to Y$ be the identity map. Then clearly $f^{-1}\colon \mathcal{O}(Y) \to \mathcal{O}(X)$ is injective, i.e. f is an epimorphism in **Loc**; but the pullback of Y_b along f is the empty sublocale of X, since it has no points and every sublocale of X is (open, by 1.1.20(ii), and hence) spatial. In particular, we note from this example that *epimorphisms in* **Loc** *are not stable under pullback.*

Lemma 1.2.13 *Let Y be a spatial locale, S a complemented sublocale of Y. Then (assuming classical logic)*

(i) *S is spatial.*

(ii) *For any morphism $f\colon X \to Y$ in* **Sob**, *the pullback of S along f in* **Loc** *coincides with the pullback in* **Sob**.

Proof (i) Again, let T denote the complement of S in $\mathrm{Sub}(Y)$. Since the terminal locale 1 has only two sublocales 1 and $\emptyset$ (this is where we use classical logic!), it is clear that any point of Y defines an atom in $\mathrm{Sub}(Y)$, and hence must factor through either $S \rightarrowtail Y$ or $T \rightarrowtail Y$. So, writing S_p for the spatial part of S as before, we have $S_p \cup T_p \geq Y_p = Y$; but $S_p \leq S$ and $T_p \leq T$, whence $S_p = S$ follows as in the proof of 1.2.11.

(ii) Clearly, a pullback of a complemented sublocale is complemented. So this follows immediately from (i) and the fact that the 'spatial part' functor $X \mapsto X_p$, being a right adjoint, preserves pullbacks. □

Not every spatial sublocale of a spatial locale is complemented. If X is a (sober) space and S a (subclosed) subspace such that both S and $X \setminus S$ are dense in X (for example, if $X = \mathbb{R}$ and $S = \mathbb{Q}$), then any sublocale T of X satisfying $S \cup T = X$ must contain all the points of $X \setminus S$; so both S and T are dense and thus $S \cap T \geq X_b$. Hence S does not have a complement in $\mathrm{Sub}(X)$. In fact it can be shown that a subclosed S defines a complemented sublocale of X iff there does *not* exist a nonempty subset $Y \subseteq X$ such that both $S \cap Y$ and $Y \setminus S$ are dense in Y; see [484].

If X is a locale, we write X_d for the locale defined by $\mathcal{O}(X_d) = N(\mathcal{O}(X))$, and call it the *dissolution* of X. Translating 1.1.19 and 1.1.20 into localic language, we see that $X \mapsto X_d$ is a functor **Loc** $\to$ **Loc**, and that there is a natural transformation $d_X\colon X_d \to X$ which is both monic and epic, and has the property that pullback along d_X induces a bijection between arbitrary sublocales of X and closed sublocales of X_d. (Thus one dissolves a locale by declaring everything to be closed.) If we assume classical logic (so that $\mathcal{O}(1)$ is Boolean), then d_X induces a bijection between the points of X_d and those of X; indeed, for a space X, the spatial part X_{dp} of X_d is simply X retopologized by declaring all subclosed sets to be closed (in particular if X is a T_D-space, this is just the discrete topology on the underlying set of X). However, X_d is rarely spatial, and the construction

$X \mapsto X_d$ is rarely idempotent. We note, on the other hand, that the morphism $d_X : X_d \to X$ is stably epic in **Loc** (that is, all its pullbacks are epic), in contrast to 1.2.12; this follows easily from 1.1.22(iii), since the bottom edge of the frame pushout displayed there is a monomorphism in **Frm**.

Next, we translate the elementary properties of fibrewise dense and fibrewise closed nuclei into localic terms. Of course, given a locale map $f: Y \to X$ and a sublocale Y' of Y, we say $Y' \rightarrowtail Y$ is fibrewise dense (resp. fibrewise closed) over X if the corresponding nucleus on $\mathcal{O}(Y)$ is fibrewise dense (resp. fibrewise closed) relative to $\mathcal{O}(X)$.

Lemma 1.2.14

(i) *Suppose given locale maps $Y' \rightarrowtail Y \to X \to X'$ of which the first is an inclusion. If $Y' \to Y$ is fibrewise dense over X (resp. fibrewise closed over X'), then it is fibrewise dense over X' (resp. fibrewise closed over X). The converse implications hold if $X \to X'$ is an inclusion.*

(ii) *If $i: Y' \rightarrowtail Y$ is fibrewise dense over X and $f: Y \to X$ is an epimorphism, then the composite $fi: Y' \to X$ is an epimorphism.*

(iii) *Arbitrary intersections of fibrewise dense sublocales (over a given base locale X) are fibrewise dense.*

(iv) *Arbitrary intersections, and finite unions, of fibrewise closed sublocales are fibrewise closed.*

(v) *Fibrewise closedness is stable under pullback; that is, given a pullback square*

$$\begin{tikzcd} Z' \arrow[r] \arrow[d, tail] & Y' \arrow[d, tail] \\ Z \arrow[r] & Y \end{tikzcd}$$

of locales over X where the vertical morphisms are inclusions, then $Z' \rightarrowtail Z$ is fibrewise closed over X if $Y' \rightarrowtail Y$ is.

(vi) *Fibrewise closure is functorial; that is, given a commutative square as in the statement of (iv), the composite $\overline{Z'}^X \rightarrowtail Z \to Y$ factors through $\overline{Y'}^X \rightarrowtail Y$, where $\overline{(-)}^X$ denotes fibrewise closure over X.*

(vii) *Fibrewise closedness is stable under change of base; that is, given a locale morphism $X' \to X$ and a fibrewise closed inclusion $Y' \rightarrowtail Y$ in $\mathbf{Loc}/X$, the induced morphism $Y' \times_X X' \rightarrowtail Y \times_X X'$ is fibrewise closed over X'.*

Proof The 'forward' implications in (i) are immediate from the definitions of fibrewise dense and fibrewise closed nuclei; the reverse ones from the fact that they depend only on the image of the frame homomorphism $\mathcal{O}(X) \to \mathcal{O}(Y)$. For (ii), we use the fact that $i_*(fi)^* = i_*i^*f^* = f^*$ if i is fibrewise dense; so $(fi)^*$ must be injective if f^* is.

(iii) is a translation of 1.1.22(i), and (iv) is a translation of the fact, included in 1.1.22(iii), that the fibrewise closed nuclei form a subframe of the frame of all nuclei. (v) also follows from 1.1.22(iii), since pulling back sublocales of Y along $g: Z \to Y$ corresponds to applying the frame homomorphism $N(g^*)$ of 1.1.19 to nuclei on $\mathcal{O}(Y)$, and it is clear from the pushout diagram in the statement of 1.1.22(iii) that this operation preserves fibrewise closed nuclei.

(vi) is now immediate from (v), since the pullback along $Z \to Y$ of the fibrewise closure of $Y' \rightarrowtail Y$ is fibrewise closed in Z and contains Z'. Finally, for (vii), we note that it follows from (v) that $Y' \times_X X' \rightarrowtail Y \times_X X'$ is fibrewise closed over X; hence by (i) it is fibrewise closed over X'. □

Unfortunately, the analogue of 1.2.14(vii) for fibrewise denseness is false in general; we shall see a counterexample shortly. However, it is true in the case when the morphism $Y \to X$ is open, as we shall see in 3.1.14 below. We note an alternative characterization of fibrewise denseness:

Lemma 1.2.15 *Let $f: Y \to X$ be a locale morphism, and $Y' \rightarrowtail Y$ a sublocale of Y. The following are equivalent:*

(i) *$Y' \rightarrowtail Y$ is fibrewise dense over X.*

(ii) *For each closed sublocale $X' \rightarrowtail X$, the pullback of $Y' \rightarrowtail Y$ along $X' \rightarrowtail X$ is a dense inclusion.*

(iii) *For each locally closed sublocale $X' \rightarrowtail X$, the pullback of $Y' \rightarrowtail Y$ along $X' \rightarrowtail X$ is a dense inclusion.*

Proof Since denseness (in the usual sense) is by definition stable under pullback along open inclusions, the equivalence of (ii) and (iii) is immediate. For (i) $\Leftrightarrow$ (ii), we note that if we pull back Y and Y' along a closed inclusion $\complement U \rightarrowtail X$, we obtain the sublocales $\complement f^*(U)$ and $\complement j(f^*(U))$ respectively (where j is the nucleus on $\mathcal{O}(Y)$ corresponding to the sublocale Y'), and so the assertion that j fixes $f^*(U)$ is exactly the assertion that the second of these is a dense sublocale of the first. □

Corollary 1.2.16 *Assuming classical logic, if X is a (sober) T_D-space, then an inclusion of spaces over X is fibrewise dense iff its pullback over every point of X is a dense inclusion.*

Proof One direction is immediate from 1.2.15 (and the fact that the points of a T_D-space, being locally closed, are complemented as sublocales, so that by 1.2.13 we do not have to worry about the difference between localic and spatial pullbacks). For the converse, we use the fact that every closed sublocale of X is spatial (1.2.6(b)), and hence the union of its points. □

The analogue of 1.2.16 for fibrewise closedness is also true, by 1.2.14(vi). Hence, given $f: Y \to X$ in **Sob**, where X is a T_D-space, and a (subclosed) subspace Y' of Y, we form its fibrewise closure over X by forming the union,

over all points $x \in X$, of the closure of $Y' \cap f^{-1}(x)$ in the fibre $f^{-1}(x)$. This explains the use of the word 'fibrewise' for these notions.

However, the T_D assumption cannot be omitted from 1.2.16. Let Y be the set of real numbers with the Euclidean topology, let X be the same set topologized so that the only nonempty open sets are the cofinite sets containing 0, let $f: Y \to X$ be the identity mapping, and let $Y' = Y \setminus \{0\}$. Then it is easily seen using the criterion of 1.2.15(ii) that $Y' \rightarrowtail Y$ is fibrewise dense over X; but its pullback over the non-locally-closed point 0 of X is not dense.

We conclude this section with some remarks on regularity and the Hausdorff property for locales. Classically, a space X is regular iff every open set can be expressed as a union of open sets whose closures it contains; correspondingly, we define a locale to be *regular* if every $U \in \mathcal{O}(X)$ satisfies $U = \bigcup\{V \in \mathcal{O}(X) \mid V \lhd U\}$, where $V \lhd U$ is an abbreviation for 'there exists $W \in \mathcal{O}(X)$ such that $V \cap W = \emptyset$ and $U \cup W = X$ (equivalently, such that $V \leq \complement W \leq U$ in $\mathrm{Sub}(X)$)'. We say X is *Hausdorff* if the diagonal $X \rightarrowtail X \times X$ is a closed sublocale: because locale products do not coincide with space products in general, this is actually a stronger condition for spaces than the usual Hausdorff axiom (that is, there exist Hausdorff spaces which are not Hausdorff as locales). For this reason, Hausdorff locales (in our sense) have been called 'strongly Hausdorff locales' by some authors (e.g. [476, 520]); however, we now prefer the unadorned name 'Hausdorff' for this property. There is no totally satisfactory analogue of the T_1 property for locales (see [886, 1057]), but the following condition is sometimes useful: we say X is a *T_U-locale* (short for 'totally unordered') if, for any two locale maps $f, g: Y \rightrightarrows X$, the relation $f \leq g$ (in the partial ordering of 1.1.6) implies $f = g$. For a (sober) space X, this clearly implies that the points of X are discretely ordered in the 'specialization' ordering ($x \leq y$ iff x is in the closure of $\{y\}$), and hence that X is a T_1-space, but once again the locale-theoretic condition is stronger (see [886]).

We stress that, in the definition of Hausdorff locale, the word 'closed' is used in its original strong sense, not in the weak 'fibrewise' sense of 1.1.22. Constructively, there are weak 'fibrewise' versions of all the basic separation axioms (see [533]), but we shall not need to consider them here.

Lemma 1.2.17

(i) *Regular locales are Hausdorff.*

(ii) *In a regular locale, every sublocale is an intersection of open sublocales.*

(iii) *Hausdorff locales are T_U.*

Proof (i) Suppose X is regular; let $(a, b): \overline{X} \rightarrowtail X \times X$ be the closure of the diagonal sublocale. We must show that the two projections a and b are equal. Now if U, V are any two opens in X such that $V \lhd U$, then since the open sublocale $\neg V \times V$ of $X \times X$ (where $\neg V$ is the complement of the closure of V, i.e. the Heyting negation of V in $\mathcal{O}(X)$) is disjoint from the diagonal it is also disjoint from $\overline{X}$; that is, $a^*(\neg V) \cap b^*(V) = \emptyset$ in $\mathcal{O}(\overline{X})$. But $\neg V \cup U$ is the whole of

X, so $a^*(\neg V) \cup a^*(U)$ is the whole of $\overline{X}$. It follows that we have $b^*(V) \leq a^*(U)$, for all such pairs (U, V). But $U = \bigcup\{V \in \mathcal{O}(X) \mid V \lhd U\}$, and b^* preserves this union, so $b^*(U) \leq a^*(U)$. By symmetry, we also have $a^*(U) \leq b^*(U)$; so $a = b$, as required.

(ii) It suffices by 1.1.17 to show that every closed sublocale is an intersection of open sublocales. But this is immediate from the definition, since for any $U \in \mathcal{O}(X)$ we have

$$\begin{aligned} \complement U &\leq \bigwedge\{\neg V \mid V \lhd U\} \\ &\leq \bigwedge\{\complement V \mid V \lhd U\} \\ &= \complement U \,. \end{aligned}$$

(iii) We claim that if $f \leq g$ then $(f, g)\colon Y \to X \times X$ factors through the closure of the diagonal. For if $U \times V$ is any open rectangle disjoint from the diagonal we have $(f,g)^*(U \times V) = f^*U \cap g^*V \leq g^*U \cap g^*V = g^*(U \cap V) = \emptyset$. Hence if the diagonal is closed we have $f = g$. □

At this point we shall break off our study of locale theory, in order to begin the development of sheaf theory. However, further parts of the theory of locales will be developed as they are required from time to time in the remaining sections of this chapter, and also in Chapters C3–C5. For a more connected account of the subject, see [520].

Suggestions for further reading: Fourman & Grayson [363], Isbell [476, 482, 484], Johnstone [520, 528, 533], Joyal & Tierney [560], Kock & Plewe [639], Plewe [980, 981], Sambin [1078, 1079], Valentini [1183, 1184], Vickers [1205].

C1.3 Sheaves, local homeomorphisms and frame-valued sets

There are three, substantially different, ways of defining the category of sheaves on a space or locale. All three of them were introduced (for spaces, at least) at different points in Part A, and they were shown to be toposes. In this section, we shall recall the three definitions (generalizing them to locales in the process) and give a direct proof that they yield equivalent categories.

The first and most familiar, originally introduced in A2.1.8, regards a sheaf as a special sort of presheaf.

Definition 1.3.1 (*a*) A *presheaf* (of sets) on a locale X is a functor $\mathcal{O}(X)^{\mathrm{op}} \to$ **Set**. Thus a presheaf F assigns to each open $U \rightarrowtail X$ a set $F(U)$ (whose elements are commonly called *sections* of F over U), and to each inclusion $V \leq U$ of open sublocales a *restriction map* $F(U) \to F(V)$, subject to the obvious compatibility properties.

(*b*) A presheaf F on X is said to be a *sheaf* if, given any family of sections $(s_i \in F(U_i) \mid i \in I)$ which are *compatible* in the sense that, for each pair (i, j), the

restrictions of s_i and s_j are equal in $F(U_i \cap U_j)$, there is a unique $s \in F(\bigcup_{i \in I} U_i)$ whose restriction to $F(U_i)$ equals s_i, for all $i \in I$. Informally, 'compatible families of sections can be uniquely patched together'.

(*c*) A *morphism of sheaves* $f: F \to G$ is simply a natural transformation of functors. We write $\mathbf{Sh}(X)$ for the category of sheaves on X.

The second definition requires consideration of local homeomorphisms between locales. Of course, we say a continuous map $p: E \to X$ of locales is a *local homeomorphism* if E can be covered by open sublocales U for which the composite $U \rightarrowtail E \to X$ is (isomorphic to) the inclusion of an open sublocale of X. We shall need the following elementary facts about local homeomorphisms:

Lemma 1.3.2

(i) *A composite of local homeomorphisms is a local homeomorphism.*

(ii) *A pullback of a local homeomorphism is a local homeomorphism.*

(iii) *Given a commutative triangle of continuous maps*

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "p"'] & & F \arrow[dl, "q"] \\ & X & \end{tikzcd} ,$$

if p and q are local homeomorphisms then so is f.

(iv) *The inclusion functor from the category of locales and local homeomorphisms to* $\mathbf{Loc}$ *creates pullbacks.*

(v) *If X is a spatial locale, then so is the domain of any local homeomorphism $p: E \to X$.*

Proof (i) Let

$$E \xrightarrow{p} X \xrightarrow{q} Y$$

be a composable pair of local homeomorphisms. Suppose that $U \in \mathcal{O}(E)$ and $V \in \mathcal{O}(X)$ are such that the composites $U \rightarrowtail E \to X$ and $V \rightarrowtail X \to Y$ are isomorphic to the inclusions of opens $\overline{U} \in \mathcal{O}(X)$, $\overline{V} \in \mathcal{O}(Y)$ respectively. Then it is easy to see that the composite $U \cap p^*(V) \rightarrowtail E \to X$ is isomorphic to $\overline{U} \cap V \rightarrowtail X$, and that this in turn composes with q to yield an open sublocale of Y. But since p^* preserves unions of open sublocales and the infinite distributive law holds in $\mathcal{O}(E)$, E can be covered by open sublocales of the form $U \cap p^*(V)$. Thus qp is a local homeomorphism.

(ii) Let

$$\begin{tikzcd} E \arrow[r, "g"] \arrow[d, "p"] & F \arrow[d, "q"] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

be a pullback square in **Loc**, where q is a local homeomorphism. For any $U \in \mathcal{O}(F)$, the square

$$\begin{tikzcd} g^*(U) \arrow[r] \arrow[d, tail] & U \arrow[d, tail] \\ E \arrow[r, "g"] & F \end{tikzcd}$$

is also a pullback by 1.2.10, and hence so is its composite with the square above; thus if U is such that $U \rightarrowtail F \to Y$ is an open inclusion, then so is $g^*(U) \rightarrowtail E \to X$. Since g^* preserves unions of open sublocales, the result follows.

(iii) It is clear that if E can be covered by open sublocales which map by local homeomorphisms to F, then $E \to F$ is itself a local homeomorphism; so we may reduce to the case where p is the inclusion of an open sublocale $U \rightarrowtail X$. Now if $V \in \mathcal{O}(F)$ is such that $V \rightarrowtail F \to X$ is the inclusion of an open sublocale $\overline{V}$, then $f^*(V)$ may be identified with the open sublocale $U \cap \overline{V}$ of U, and the composite $U \cap \overline{V} \rightarrowtail U \to F$ with the inclusion $q^*(U) \cap V \rightarrowtail F$. So, once again, the result follows from the fact that f^* preserves unions of open sublocales.

(iv) From (ii), we know that if $p\colon E \to X$ and $q\colon F \to X$ are local homeomorphisms, then so are the projections from the pullback $E \times_X F$ in **Loc**. This does not suffice to show that it is a pullback in the subcategory; we need to show that if $f\colon Z \to E$ and $g\colon Z \to F$ are local homeomorphisms satisfying $pf = qg$, then the induced map $(f, g)\colon Z \to E \times_X F$ is also a local homeomorphism. But this follows from (iii), since (by (i)) both Z and $E \times_X F$ admit local homeomorphisms into X.

(v) If X is spatial, then so is any open sublocale of X (1.2.6(a)); hence the assertion that $p\colon E \to X$ is a local homeomorphism implies that E can be covered by spatial (open) sublocales, and is thus spatial. □

In this chapter, we shall write **LH** for the category of locales and local homeomorphisms between them. In Part A, we used this name for the category of spaces and local homeomorphisms, but the difference is not a substantial one: although we have 'enlarged' the category **LH**, Lemma 1.3.2(v) tells us that we have not changed the category **LH**$/X$ when X is a space. 1.3.2(iii) says that **LH**$/X$ is a full subcategory of **Loc**$/X$, for any X; and 1.3.2(iv), plus the (obvious) fact that identity morphisms are local homeomorphisms, says that it is

closed under finite limits. Our second definition of the category of sheaves on X is simply as the slice category $\mathbf{LH}/X$ (cf. A1.2.7).

We remark in passing that local homeomorphisms are often known as 'étale maps'; this is based on a mistranslation of the French term 'espace étalé' (literally, 'displayed space') for (the domain of) the local homeomorphism over X corresponding to a given sheaf on X under the equivalence of 1.3.11 below, and does not derive from the 'étale morphisms' of algebraic geometry. For this reason, we shall avoid using the term as far as possible.

The third definition of the category of sheaves is as the category of $\mathcal{O}(X)$-valued sets, which was introduced at the end of Section A3.3 as the category of maps in a certain allegory $\mathbf{Mat}(\mathcal{O}(X))[\check{S}]$. We give a more explicit definition here:

Definition 1.3.3 (*a*) An *$\mathcal{O}(X)$-valued set* is a pair (A, ϵ) where A is a set and $\epsilon \colon A \times A \to \mathcal{O}(X)$ is a function satisfying

$$\epsilon(a_1, a_2) = \epsilon(a_2, a_1)$$

and

$$\epsilon(a_1, a_2) \cap \epsilon(a_2, a_3) \leq \epsilon(a_1, a_3)$$

for all $a_1, a_2, a_3 \in A$. Thus ϵ is a 'symmetric and transitive $\mathcal{O}(X)$-valued relation on A'. We do not require ϵ to be reflexive; we sometimes abbreviate $\epsilon(a, a)$ to $\epsilon(a)$, and call it the *extent* of the element a.

(*b*) A *map of $\mathcal{O}(X)$-valued sets* $\phi \colon (A, \epsilon) \to (B, \delta)$ is a function $\phi \colon B \times A \to \mathcal{O}(X)$ satisfying

$$\begin{gathered} \phi(b, a) \leq \delta(b) \cap \epsilon(a), \\ \phi(b_1, a_1) \cap \delta(b_1, b_2) \cap \epsilon(a_1, a_2) \leq \phi(b_2, a_2), \\ \phi(b_1, a) \cap \phi(b_2, a) \leq \delta(b_1, b_2), \end{gathered}$$

and

$$\epsilon(a) \leq \bigcup \{\phi(b, a) \mid b \in B\}$$

for all $a, a_1, a_2 \in A$ and all $b, b_1, b_2 \in B$. Informally, ϕ is an '$\mathcal{O}(X)$-valued functional relation from (A, ϵ) to (B, δ)'.

(*c*) Given maps $\phi \colon (A, \epsilon) \to (B, \delta)$ and $\psi \colon (B, \delta) \to (C, \gamma)$, the composite $\psi\phi \colon (A, \epsilon) \to (C, \gamma)$ is defined by

$$\psi\phi(c, a) = \bigcup \{\psi(c, b) \cap \phi(b, a) \mid b \in B\} .$$

It is straightforward to verify that $\psi\phi$ is a map of $\mathcal{O}(X)$-valued sets (i.e. that it satisfies the conditions in (*b*)), and that composition is associative. Further, ϵ itself acts as an identity map $(A, \epsilon) \to (A, \epsilon)$ for any $\mathcal{O}(X)$-valued set (A, ϵ); so we have a category of $\mathcal{O}(X)$-valued sets, which we denote by $\mathbf{Set}(\mathcal{O}(X))$.

We now embark on proving the equivalence of the categories $\mathbf{Sh}(X)$, $\mathbf{LH}/X$ and $\mathbf{Set}(\mathcal{O}(X))$, for any locale X. First we define a functor $\Gamma\colon \mathbf{Loc}/X \to \mathbf{Sh}(X)$: given a locale $p\colon E \to X$ over X, $\Gamma(p)$ is the presheaf whose sections over $U \in \mathcal{O}(X)$ are the continuous sections of p over U, i.e. the commutative triangles

$$\begin{tikzcd} & E \arrow[dr, "p"] & \\ U \arrow[ur, "s"] \arrow[rr, tail] & & X \end{tikzcd}$$

in **Loc**. To verify that $\Gamma(p)$ is a sheaf, we need to show that open coverings are colimits in **Loc**, i.e.

Lemma 1.3.4 *Let X be a locale, and $(U_i \mid i \in I)$ a family of open sublocales of X with union U. Then the inclusion maps $U_i \rightarrowtail U$ form a colimit cone for the diagram of locales whose vertices are all the U_i and the $U_i \cap U_j$, and whose edges are the inclusions $U_i \cap U_j \rightarrowtail U_i$ and $U_i \cap U_j \rightarrowtail U_j$.*

Proof We recall that colimits in **Loc** are just limits in **Frm**, and that the latter are created by the forgetful functor $\mathbf{Frm} \to \mathbf{Set}$. So this is easy to verify, bearing in mind that $\mathcal{O}(U)$ may be identified with the principal ideal $\downarrow(U)$ in $\mathcal{O}(X)$, and the frame map $\mathcal{O}(U) \to \mathcal{O}(U_i)$ corresponding to the inclusion is identified with $(-) \cap U_i$. □

It follows that a compatible family of continuous maps $s_i\colon U_i \to E$ may be uniquely patched together to a continuous map $s\colon U \to E$; and if the s_i are all sections of p, then so is s, by the uniqueness of maps out of colimits. Thus $\Gamma(p)$ is a sheaf, as claimed. Also, it is clear that if

$$\begin{tikzcd} E \arrow[rr, "f"] \arrow[dr, "p"'] & & F \arrow[dl, "q"] \\ & X & \end{tikzcd}$$

is a map of locales over X, then composing sections of p with f yields a morphism of sheaves $\Gamma(f)\colon \Gamma(p) \to \Gamma(q)$. Thus Γ becomes a functor $\mathbf{Loc}/X \to \mathbf{Sh}(X)$. (Its restriction to the subcategory $\mathbf{LH}/X$ will be part of the equivalence which we seek.)

Next, we define a functor $\Theta\colon [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}(\mathcal{O}(X))$. Given two sections $s \in F(U)$, $t \in F(V)$ of a presheaf F on X, we let $\epsilon_F(s,t)$ denote the union of all open $W \leq U \cap V$ such that the restrictions of s and t to W are equal. (If F is a sheaf, then we may alternatively define $\epsilon_F(s,t)$ as the largest open sublocale

on which the restrictions of s and t are equal, because we can patch together their restrictions to all the open sublocales where they agree.) In particular, $\epsilon_F(s) = \epsilon_F(s,s)$ is just the open U such that $s \in F(U)$. It is now easy to see that the disjoint union $\Theta(F) = \coprod_{U \in \mathcal{O}(X)} F(U)$, structured with the map $\epsilon_f \colon \Theta(F) \times \Theta(F) \to \mathcal{O}(X)$, is an $\mathcal{O}(X)$-valued set, as defined in 1.3.3(a). Given a morphism of presheaves $f \colon F \to G$, we define $\Theta(f) \colon \Theta(G) \times \Theta(F) \to \mathcal{O}(X)$ by

$$\Theta(f)(t,s) = \epsilon_G(t, f(s)) .$$

Again, it is straightforward to verify that $\Theta(f)$ is a map of $\mathcal{O}(X)$-valued sets; the key observation is that we have $\epsilon_F(s_1,s_2) \leq \epsilon_G(f(s_1), f(s_2))$ for any $s_1, s_2 \in \Theta(F)$. And if $g \colon G \to H$ is another morphism of presheaves, then

$$\begin{aligned}\Theta(g)\Theta(f)(u,s) &= \bigcup\{\epsilon_H(u,g(t)) \cap \epsilon_G(t,f(s)) \mid t \in \Theta(G)\} \\ &\geq \epsilon_H(u,gf(s)) \cap \epsilon_G(f(s),f(s)) \\ &= \epsilon_H(u,gf(s)) \cap \epsilon_H(gf(s),gf(s)) \\ &= \epsilon_H(u,gf(s))\end{aligned}$$

and conversely

$$\epsilon_H(u,g(t)) \cap \epsilon_G(t,f(s)) \leq \epsilon_H(u,g(t)) \cap \epsilon_H(g(t),gf(s)) \leq \epsilon_H(u,gf(s))$$

for any t; so $\Theta(g)\Theta(f) = \Theta(gf)$, and we have proved

Lemma 1.3.5 *Θ is a functor $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}(\mathcal{O}(X))$.* □

The second of our circle of three functors will be the restriction of Θ to the subcategory $\mathbf{Sh}(X)$. Before we complete the circle by defining a functor $\mathbf{Set}(\mathcal{O}(X)) \to \mathbf{LH}/X$, we pause to note that the $\mathcal{O}(X)$-valued sets which are the images of sheaves under the functor Θ are of a rather special type. Given an $\mathcal{O}(X)$-set (A,ϵ), we define a *singleton* of (A,ϵ) to be a mapping $\sigma \colon A \to \mathcal{O}(X)$ satisfying

(i) $\sigma(a) \cap \epsilon(a,b) \leq \sigma(b)$ and
(ii) $\sigma(a) \cap \sigma(b) \leq \epsilon(a,b)$

for all $a, b \in A$. For any $a \in A$, the mapping $b \mapsto \epsilon(a,b)$ is easily seen to be a singleton, which we denote by $\tilde{a}$.

Definition 1.3.6 We say an $\mathcal{O}(X)$-set (A,ϵ) is *complete* if every singleton of (A,ϵ) is of the form $\tilde{a}$ for a unique $a \in A$.

Lemma 1.3.7 *For any sheaf F on X, the $\mathcal{O}(X)$-valued set $\Theta(F)$ is complete.*

Proof Let σ be a singleton of $\Theta(F)$, and consider the family of all sections $s_i \in F(U_i)$ such that $\sigma(s_i) = U_i$. From the definition of a singleton, it is clear

that each $F(U_i)$ contains at most one such element, and that the set of all such elements is closed under restriction; so they form a compatible family. Now it is straightforward to verify that $\sigma = \tilde{s}$ iff s is the unique section obtained by patching the s_i together. $\square$

In the converse direction, if $a \in A$ and $U \leq \epsilon(a)$, then the function $b \mapsto \epsilon(a,b) \cap U$ is readily seen to be a singleton; so the assertion that (A, ϵ) is complete implies that for every such pair (a, U) there exists a unique $a|_U \in A$ with $\epsilon(a|_U) = \epsilon(a, a|_U) = U$. Thus the assignment $U \mapsto \{a \in A \mid \epsilon(a) = U\}$ can be made into a presheaf F on X. And this presheaf is a sheaf, because if $(a_i \in F(U_i) \mid i \in I)$ is a compatible family of sections of it, then the function $b \mapsto \bigcup_{i \in I} \epsilon(a_i, b)$ is again a singleton, and it is of the form $\tilde{a}$ iff a is a patching of the a_i. Thus the $\mathcal{O}(X)$-valued sets which occur in the image of Θ are precisely the complete ones.

Complete $\mathcal{O}(X)$-valued sets are easier to work with than arbitrary ones, because maps between them have a particularly simple description:

Lemma 1.3.8 *Let $\phi\colon (A, \epsilon) \to (B, \delta)$ be a map of $\mathcal{O}(X)$-valued sets, and suppose (B, δ) is complete. Then there is a unique function $f\colon A \to B$ such that $\phi(b, a) = \delta(b, f(a))$ for all $a \in A$, $b \in B$. Moreover, we have $\delta(f(a_1), f(a_2)) \geq \epsilon(a_1, a_2)$ for all $a_1, a_2 \in A$, with equality if $a_1 = a_2$.*

Proof For each $a \in A$, the mapping $b \mapsto \phi(b, a)$ is a singleton of (B, δ), so we can write it as $\widetilde{f(a)}$ for a unique $f(a)$. This defines the function f. Now we have

$$\delta(f(a)) = \bigcup_{b \in B} \delta(b, f(a)) = \bigcup_{b \in B} \phi(b, a) = \epsilon(a)$$

by one of the identities in the definition of a map of $\mathcal{O}(X)$-valued sets, which proves the last assertion of the lemma; and

$$\begin{aligned}
\delta(f(a_1), f(a_2)) &= \phi(f(a_1), a_2) \\
&\geq \phi(f(a_1), a_1) \cap \epsilon(a_1, a_2) \\
&= \delta(f(a_1)) \cap \epsilon(a_1, a_2) \\
&= \epsilon(a_1) \cap \epsilon(a_1, a_2) = \epsilon(a_1, a_2)
\end{aligned}$$

which proves the penultimate one. $\square$

It is also easy to verify that composition of maps between complete $\mathcal{O}(X)$-valued sets, as defined in 1.3.3(*c*), corresponds to ordinary composition of the representing functions as defined in 1.3.8; so the full subcategory of complete $\mathcal{O}(X)$-valued sets becomes a concrete category. Moreover, we do not lose anything by restricting our attention to complete $\mathcal{O}(X)$-valued sets:

Lemma 1.3.9 *For any $\mathcal{O}(X)$-valued set (A, ϵ), there is a complete $\mathcal{O}(X)$-valued set $(\tilde{A}, \tilde{\epsilon})$ which is isomorphic to it in* $\mathbf{Set}(\mathcal{O}(X))$.

Proof We define $\widetilde{A}$ to be the set of all singletons of (A, ϵ), with

$$\tilde{\epsilon}(\sigma, \tau) = \bigcup_{a \in A} (\sigma(a) \cap \tau(a)) .$$

It is straightforward to verify that $(\widetilde{A}, \widetilde{\epsilon})$ is an $\mathcal{O}(X)$-valued set. Moreover, if $a, b \in A$, then

$$\begin{aligned}\tilde{\epsilon}(\tilde{a}, \tilde{b}) &\geq \tilde{a}(a) \cap \tilde{a}(b) \\ &= \epsilon(a) \cap \epsilon(a, b) = \epsilon(a, b),\end{aligned}$$

the reverse inequality being trivial. Now if τ is a singleton of $(\tilde{A}, \tilde{\epsilon})$, let $\sigma\colon A \to \mathcal{O}(X)$ be defined by $\sigma(a) = \tau(\tilde{a})$; then it is easily verified that σ is a singleton of (A, ϵ), and it is the unique singleton such that $\tau = \tilde{\sigma}$. Thus $(\tilde{A}, \tilde{\epsilon})$ is complete. Finally, we define $\iota\colon \tilde{A} \times A \to \mathcal{O}(X)$ by $\iota(\sigma, a) = \sigma(a)$; then ι satisfies the conditions of 1.3.3(b) both as a map $(A, \epsilon) \to (\tilde{A}, \tilde{\epsilon})$ and as a map $(\tilde{A}, \tilde{\epsilon}) \to (A, \epsilon)$, and the composites of these two maps either way round are the appropriate identity maps. □

Of course, the function $A \to \tilde{A}$ which corresponds, via 1.3.8, to the isomorphism $\iota\colon (A, \epsilon) \to (\tilde{A}, \tilde{\epsilon})$ is simply $a \mapsto \tilde{a}$. With a little more effort, one may show that the construction $(A, \epsilon) \mapsto (\tilde{A}, \tilde{\epsilon})$ of 1.3.9 is functorial, and that ι is a natural isomorphism; thus the category $\mathbf{Set}(\mathcal{O}(X))$ is equivalent to its full subcategory $\mathbf{CSet}(\mathcal{O}(X))$ of complete $\mathcal{O}(X)$-valued sets. (Combining this with the remarks after 1.3.7, it is now easy to construct an inverse equivalence for $\Theta\colon \mathbf{Sh}(X) \to \mathbf{Set}(\mathcal{O}(X))$.) For the next step of constructing a functor to $\mathbf{LH}/X$, we shall take our domain to be $\mathbf{CSet}(\mathcal{O}(X))$ rather than $\mathbf{Set}(\mathcal{O}(X))$; this is not strictly essential, but it does considerably simplify the description of how the functor acts on morphisms.

Given a (complete) $\mathcal{O}(X)$-valued set (A, ϵ), we construct a diagram $D(A, \epsilon)$ in $\mathbf{Loc}/X$ as follows. For the vertices, we take one copy of the inclusion $\epsilon(a) \rightarrowtail X$ for each $a \in A$, and one copy of $\epsilon(a, b) \rightarrowtail X$ for each pair (a, b). And for each (a, b), we take the inclusions $\epsilon(a, b) \rightarrowtail \epsilon(a)$ and $\epsilon(a, b) \rightarrowtail \epsilon(b)$ as edges of the diagram. We define $\Delta(A, \epsilon)$ to be the colimit of this diagram in $\mathbf{Loc}/X$; for the moment, let us denote it by $p\colon E \to X$.

We need to prove that p is a local homeomorphism: but for each $a \in A$, the ath leg $\lambda_a\colon \epsilon(a) \to E$ of the colimiting cone is a sublocale inclusion (i.e. a regular monomorphism), since its composite with p is an (open) inclusion. And these sublocales clearly cover E; so it suffices to prove that they are open. But since $\mathcal{O}(E)$ is a limit in $\mathbf{Frm}$, we can give an explicit description of it, much as in 1.3.4: its elements correspond to families $(U_a \mid a \in A)$ of opens in X such that $U_a \leq \epsilon(a)$ for each a and $U_a \cap \epsilon(a, b) = U_b \cap \epsilon(a, b)$ for each (a, b). Under this correspondence, (the image of) λ_a clearly corresponds to the family $(\epsilon(a, b) \mid b \in A)$.

So far, we have not made any use of completeness of (A, ϵ); but in defining the effect of Δ on morphisms of $\mathcal{O}(X)$-valued sets, it will be convenient to assume that they are induced by functions on the underlying sets as in 1.3.8. So let $\phi\colon (A, \epsilon) \to (B, \delta)$ be a map of $\mathcal{O}(X)$-valued sets, and let $f\colon A \to B$ be the function corresponding to it. Then each vertex of the form $\epsilon(a)$ in the diagram $D(A, \epsilon)$ maps by the identity to the vertex $\delta(f(a))$ of $D(B, \delta)$, and each vertex $\epsilon(a, b)$ maps by an inclusion to $\delta(f(a), f(b))$; composing these maps with the appropriate legs of the colimit cone for $D(B, \delta)$, we obtain a cone under $D(A, \epsilon)$, and hence a continuous map (necessarily a local homeomorphism, by 1.3.2(iii)) from the colimit $\Delta(A, \epsilon)$ to $\Delta(B, \delta)$. It is easy to verify that this construction is functorial, and so we have

Lemma 1.3.10 Δ *is a functor* $\mathbf{CSet}(\mathcal{O}(X)) \to \mathbf{LH}/X$. $\square$

Theorem 1.3.11 *For any locale X, the categories* $\mathbf{Sh}(X)$, $\mathbf{LH}/X$, $\mathbf{Set}(\mathcal{O}(X))$ *and* $\mathbf{CSet}(\mathcal{O}(X))$ *are all equivalent.*

Proof The equivalence of $\mathbf{Set}(\mathcal{O}(X))$ and $\mathbf{CSet}(\mathcal{O}(X))$ was dealt with in 1.3.9 and the remarks immediately following it. And we have constructed functors

$$\mathbf{Sh}(X) \xrightarrow{\Theta} \mathbf{CSet}(\mathcal{O}(X)) \xrightarrow{\Delta} \mathbf{LH}/X \xrightarrow{\Gamma} \mathbf{Sh}(X);$$

so we need to show that the composite of these three functors, in any of the three possible orders, is isomorphic to the identity.

First, let F be a sheaf on X. Each $s \in F(U)$ can be regarded as an element $s \in \Theta(F)$ with $\epsilon(s) = U$, and so it defines a vertex of the diagram $D(\Theta(F))$ which is a copy of the inclusion $U \rightarrowtail X$. Thus the sth leg λ_s of the colimiting cone is a continuous section of $\Delta\Theta(F)$ over U. Moreover, if t is the restriction of s to an open $V \leq U$, then $\epsilon(s, t) = V$, and so the commutative diagram

$$\begin{tikzcd} \epsilon(s,t) \arrow[r, tail] \arrow[d, tail] & \epsilon(s) \arrow[d, "\lambda_s"] \\ \epsilon(t) \arrow[r, "\lambda_t"] & \Delta\Theta(F) \end{tikzcd}$$

implies that λ_t is the restriction of λ_s to V; thus $s \mapsto \lambda_s$ is a natural transformation, i.e. a morphism of sheaves $\lambda\colon F \to \Gamma\Delta\Theta(F)$. And from the definition of the three functors on morphisms, it is readily seen that λ is itself natural in F.

So it remains to check that λ is bijective. Given an open $U \rightarrowtail X$ and a continuous section s of $\Delta\Theta(F)$ over U, we may form the pullback

$$\begin{tikzcd} V_t \arrow[r, tail] \arrow[d, tail] & V \arrow[d, "\lambda_t"] \\ U \arrow[r, "s"] & \Delta\Theta(F) \end{tikzcd}$$

for each $t \in F(V)$ (as V varies over $\mathcal{O}(X)$). Note that λ_t and s are both (local homeomorphisms, and hence) open inclusions; so V_t is an element of $\mathcal{O}(X)$ contained in $U \cap V$. And since the λ_t, as t varies over all sections of F, cover $\Delta\Theta(F)$, the V_t form an open covering of U. Moreover, the restrictions $t|_{V_t}$ form a compatible family, since $V_t \cap V_{t'} \leq \epsilon(t, t')$, and so they patch to a unique section $\overline{s} \in F(U)$, which is easily seen to be the unique element of $F(U)$ satisfying $\lambda_{\overline{s}} = s$. Thus λ is a natural isomorphism from the identity functor on $\mathbf{Sh}(X)$ to $\Gamma\Delta\Theta$.

The proof that $\Theta\Gamma\Delta$ is isomorphic to the identity on $\mathbf{CSet}(\mathcal{O}(X))$ is almost identical to the above; recall that we have already, after 1.3.7, observed that sheaves are essentially the same thing as complete $\mathcal{O}(X)$-valued sets.

Finally, we consider the composite $\Delta\Theta\Gamma$. Given a local homeomorphism $p: E \to X$, the elements of $\Theta\Gamma(p)$ are the continuous sections of p over arbitrary open sublocales of X, and they define a cone (with vertex p) under the diagram $D(\Theta\Gamma(p))$ in $\mathbf{Loc}/X$, and thus a morphism $\phi_p: \Delta\Theta\Gamma(p) \to p$ in $\mathbf{Loc}/X$ (and hence in $\mathbf{LH}/X$). Once again, it is easy to see that ϕ_p is natural in p, so we have only to verify that it is an isomorphism for all p.

Let s be a continuous section of p over $U \in \mathcal{O}(X)$. Then the diagram

$$\begin{tikzcd} U \arrow[r, equal] \arrow[d, "\lambda_s"] & U \arrow[d, "s"] \\ \Delta\Theta\Gamma(E) \arrow[r, "\phi_p"] & E \end{tikzcd}$$

(where we have written $\Delta\Theta\Gamma(E)$ for the domain of $\Delta\Theta\Gamma(p)$) commutes, by the definition of ϕ_p; but since $\Gamma(p)$ is a sheaf, it is not hard to verify that this square is a pullback. Noting that the vertical morphisms in the diagram are in fact open inclusions, we may rephrase this as saying that, for each $U \in \mathcal{O}(E)$ such that $U \rightarrowtail E \to X$ is an open inclusion, the pullback of ϕ_p along $U \rightarrowtail E$ is an isomorphism. But by assumption E is covered by such opens U, and hence ϕ_p is an isomorphism. $\square$

We recall that two of the three functors used in the proof of 1.3.11 were originally defined on larger categories than those on which we used them. (Of course,

the extension of Θ to $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ does not take values in $\mathbf{CSet}(\mathcal{O}(X))$; but we can remedy this by composing it with the completion functor of 1.3.9.) If we analyse how much of the proof survives without the assumption that F is a sheaf (resp. that $p\colon E \to X$ is a local homeomorphism), we see that we still have a natural map $\lambda\colon F \to \Gamma\Delta\Theta(F)$ (resp. $\phi\colon \Delta\Theta\Gamma(p) \to p$), which becomes an isomorphism whenever the additional hypothesis is satisfied. Thus we have

Scholium 1.3.12

(i) *The inclusion* $\mathbf{Sh}(X) \to [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ *has a left adjoint, given by the composite* $\Gamma\Delta\Theta$.

(ii) *The inclusion* $\mathbf{LH}/X \to \mathbf{Loc}/X$ *has a right adjoint, given by the composite* $\Delta\Theta\Gamma$. □

Of course, the left adjoint to the inclusion $\mathbf{Sh}(X) \to [\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ may be constructed by other means; see for example A4.4.4, or 2.2.6 below. We note that either half of 1.3.12 implies that $\mathbf{Sh}(X)$ is complete and cocomplete, since both $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$ and $\mathbf{Loc}/X$ are complete and cocomplete.

Much of the power of sheaf theory lies in the fruitful interplay of the three different, but equivalent, descriptions of the category of sheaves on a locale: whenever we need to prove something about categories of sheaves, we can appeal to whichever description yields the simplest proof. For example,

Corollary 1.3.13 *Any slice category of a category of the form* $\mathbf{Sh}(X)$ *is again of this form.*

Proof This is easy if we regard sheaves as local homeomorphisms, because (as we saw in Section A1.1) $(\mathbf{LH}/X)/(p\colon E \to X)$ is isomorphic to $\mathbf{LH}/E$. □

In other cases, it may be equally easy to use any of the three descriptions:

Proposition 1.3.14 *The equivalent categories of* 1.3.11 *are all toposes.*

Proof Of course, it suffices to prove that any one of them is a topos. But we have already given three different proofs of this in Part A: for $\mathbf{Sh}(X)$ in A2.1.8 (see also A4.3.9), for $\mathbf{Set}(\mathcal{O}(X))$ in A3.4.13, and for $\mathbf{LH}/X$ in A4.2.4(e). (The last of these three relied on the fact that X was a space rather than a general locale, but the other two did not.) We shall not repeat the details here. □

To conclude this section, we note that the locale X may be recovered, up to isomorphism, from its category of sheaves:

Proposition 1.3.15 *For any locale X, the lattice of subterminal objects (i.e. subobjects of* 1*) in* $\mathbf{Sh}(X)$ *is isomorphic to* $\mathcal{O}(X)$.

Proof It is easy to see (it follows from the fact that $\mathbf{Sh}(X)$ is reflective in $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$) that the monomorphisms in $\mathbf{Sh}(X)$ are exactly the injective natural transformations between sheaves. Since the terminal object of $\mathbf{Sh}(X)$ is the constant functor with value 1, a sub-presheaf S of it is thus determined by the

set $\{U \in \mathcal{O}(X) \mid S(U) = 1\}$, which must be a downward-closed subset of $\mathcal{O}(X)$. But if S is a sheaf, this subset must be closed under coverings, and is thus a principal ideal $\downarrow(U)$, for some $U \in \mathcal{O}(X)$. This establishes a bijection between $\mathrm{Sub}(1)$ and $\mathcal{O}(X)$, which is readily seen to be order-preserving.

Alternatively, we could argue in terms of local homeomorphisms. It is clear that the inclusion of an open sublocale is a local homeomorphism, and it is again not difficult to verify that these are exactly the monomorphisms in **LH**. Thus the subterminal objects of $\mathbf{LH}/X$ (equivalently, the subobjects of X in **LH**) are exactly the (inclusions of) open sublocales of X. □

Suggestions for further reading: Fourman & Scott [366], Gray [415], Tennison [1163].

C1.4 Continuous maps

In the previous section, we considered sheaves on a single locale X. We are now ready to investigate the functoriality of the construction $X \mapsto \mathbf{Sh}(X)$.

Let $f: X \to Y$ be a continuous map of locales. By 1.3.2(ii), we know the pullback functor $f^*: \mathbf{Loc}/Y \to \mathbf{Loc}/X$ restricts to a functor (which we shall still denote by f^*) from $\mathbf{LH}/Y$ to $\mathbf{LH}/X$, and by 1.3.11 we can regard this as a functor $f^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$, which we call the *inverse image functor* associated with f.

Lemma 1.4.1 *For any f, $f^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ is cartesian.*

Proof As a functor $\mathbf{Loc}/Y \to \mathbf{Loc}/X$, f^* preserves all limits, since it has a left adjoint (A1.2.8). And the inclusion $\mathbf{LH}/X \to \mathbf{Loc}/X$ creates finite limits, by 1.3.2(iv); so the restriction of f^* preserves them. □

Of course, we have also used f^* as the name for the frame homomorphism $\mathcal{O}(Y) \to \mathcal{O}(X)$ corresponding to f. This notation is consistent with that just introduced: if we identify open sublocales of Y with subterminal objects of $\mathbf{LH}/Y$, as in 1.3.15, then 1.2.10 tells us that $f^*(U)$ is indeed the pullback of U along f.

Remark 1.4.2 Before moving on, we pause to note a special case of the construction of 1.4.1. Suppose X is the terminal locale 1, and Y is a (sober) space; then continuous maps $X \to Y$ may be identified with points of Y. Moreover, the domain of any local homeomorphism $p: E \to Y$ is a space, by 1.3.2(v). Now the inclusion $\mathbf{Sob} \to \mathbf{Loc}$ does not preserve all pullbacks, as we saw in Section C1.2; but by 1.2.10 it preserves pullbacks of open inclusions, and from this one may easily verify that it preserves pullbacks of local homeomorphisms (essentially, because the proof of 1.3.2(ii) works in spaces as well as locales). Thus, for a point $y: 1 \to Y$, the functor $y^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(1) \simeq \mathbf{Set}$ sends $p: E \to Y$ to the fibre $E_y = \{e \in E \mid p(e) = y\}$ of p over y (which of course has the discrete

topology as a subspace of E). We call this set the *stalk* of p (or of the sheaf on Y which corresponds to it) at the point y.

Next, we observe that $f^*: \mathcal{O}(Y) \to \mathcal{O}(X)$ preserves covers and finite intersections – that is, those operations on open sublocales which are used in the definition of a sheaf (1.3.1). Thus it is immediate that if F is a sheaf on X, the composite

$$\mathcal{O}(Y)^{\mathrm{op}} \xrightarrow{f^*} \mathcal{O}(X)^{\mathrm{op}} \xrightarrow{F} \mathbf{Set}$$

is a sheaf on Y. And this construction is clearly functorial in F, so it defines a functor $f_*: \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, which we call the *direct image functor* associated with f.

Theorem 1.4.3 *For any $f: X \to Y$, the inverse image functor $f^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ is left adjoint to $f_*: \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$.*

Proof Let $p: E \to Y$ be a local homeomorphism, and F a sheaf on X. By the proof of 1.3.11, we know that p is a colimit in $\mathbf{Loc}/Y$ of a diagram $D(\Theta\Gamma(p))$ whose vertices are open inclusions $U_i \rightarrowtail Y$, the legs $\lambda_i: U_i \to E$ of the colimiting cone themselves being open inclusions. Since f^* preserves unions of open sublocales, it follows that $f^*(p)$ is the union of the open sublocales $f^*(\lambda_i): f^*(U_i) \rightarrowtail f^*(E)$, and hence (by 1.3.4) it is the colimit in $\mathbf{Loc}/X$ of the $f^*(U_i)$. Thus a morphism $f^*(p) \to \Delta\Theta(F)$ in $\mathbf{LH}/X$ corresponds to a compatible family of morphisms $U_i \to \Delta\Theta(F)$, i.e. to a compatible family of sections $s_i \in \Gamma\Delta\Theta(F)(f^*(U_i)) \cong F(f^*(U_i))$. (Warning: 'compatible' here means that the restrictions of s_i and s_j agree on the open sublocale which is the intersection of $f^*(U_i)$ and $f^*(U_j)$ in $f^*(E)$, not on their intersection in X.) But precisely the same data defines a morphism $p \to \Delta\Theta(f_*(F))$ in $\mathbf{LH}/Y$; so we have set up a bijection between these two hom-sets, which is clearly natural in both p and F. $\square$

Combining 1.4.1 and 1.4.3, we see that the pair of functors (f_*, f^*) forms a geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, as defined in A4.1.1. (In A4.1.11 we gave an alternative construction of this morphism, in which both f_* and f^* were defined in terms of sheaves rather than local homeomorphisms; of course, by the uniqueness of adjoints, the functor f^* defined there is naturally isomorphic to the one introduced in this section.)

If $g: Y \to Z$ is a continuous map of locales composable with f, then the composite frame homomorphism $(gf)^*: \mathcal{O}(Z) \to \mathcal{O}(X)$ is simply f^*g^*, from which it follows easily that the functor $(gf)_*: \mathbf{Sh}(X) \to \mathbf{Sh}(Z)$ is the composite g_*f_*. Similarly, it is easy to see that $(gf)^*: \mathbf{LH}/Z \to \mathbf{LH}/X$ is canonically isomorphic (though not necessarily equal) to the composite f^*g^*. Thus we have

Proposition 1.4.4 *The assignment $X \mapsto \mathbf{Sh}(X)$ is the object-map of a (pseudo-)functor $\mathbf{Loc} \to \mathfrak{Top}$.* $\square$

But there is a further level of structure to be considered. Recall from 1.1.6 that **Loc** is actually a 2-category (albeit a locally ordered one; i.e. its hom-categories are posets): given a parallel pair of continuous maps $f, g \colon X \rightrightarrows Y$, we write $f \leq g$ if $f^*(V) \leq g^*(V)$ for all $V \in \mathcal{O}(Y)$. (On the subcategory **Sob**, this structure is normally defined in terms of the 'specialization' partial ordering on points: given points x, y of a space X, we write $x \leq y$ if every open set containing x also contains y – note that the T_0 axiom is equivalent to the assertion that this relation is a partial order rather than a preorder, but we are working under the assumption that all spaces are sober and hence T_0 – and then, given continuous maps of spaces $f, g \colon X \rightrightarrows Y$, we set $f \leq g$ if $f(x) \leq g(x)$ for all $x \in X$. It is not hard to verify that this definition is equivalent to saying that $f^{-1}(V) \subseteq g^{-1}(V)$ for all $V \in \mathcal{O}(Y)$.)

Now we may regard an inequality $f \leq g$ as a natural transformation $f^* \to g^*$, when the latter are regarded as functors between the posets $\mathcal{O}(Y)$ and $\mathcal{O}(X)$; and if we regard them as functors $\mathcal{O}(Y)^{\mathrm{op}} \rightrightarrows \mathcal{O}(X)^{\mathrm{op}}$, then the direction of the natural transformation is of course reversed. Thus we see that the inequality $f \leq g$ induces a natural transformation $g_* \to f_* \colon \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, and hence by adjointness a natural transformation $f^* \to g^* \colon \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$. Alternatively, we could have constructed the latter directly: if $p \colon E \to Y$ is a local homeomorphism, then for each $V \in \mathcal{O}(E)$ such that $V \rightarrowtail E \to Y$ is an open inclusion, we have a morphism $f^*(V) \rightarrowtail g^*(V)$ in $\mathbf{Loc}/X$, and these morphisms combine to give a natural transformation between the diagrams in $\mathbf{Loc}/X$ whose colimits are $f^*(E)$ and $g^*(E)$ respectively. It is clear that this construction yields a functor from the ordered set $\mathbf{Loc}\,(X, Y)$ to the category $\mathfrak{Top}\,(\mathbf{Sh}(X), \mathbf{Sh}(Y))$ of geometric morphisms and geometric transformations between them; so the assignment $X \mapsto \mathbf{Sh}(X)$ is actually a 2-functor. Moreover, we have

Proposition 1.4.5 *Let X and Y be locales. Then the functor* $\mathbf{Loc}\,(X, Y) \to \mathfrak{Top}\,(\mathbf{Sh}(X), \mathbf{Sh}(Y))$ *just described is (one half of) an equivalence of categories.*

Proof Let us, for the moment, write $G(f)$ for the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by a continuous map $f \colon X \to Y$. If $g \colon \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ is an arbitrary geometric morphism, then since g^* preserves 1 and monomorphisms it restricts to a functor $\mathrm{Sub}_{\mathbf{Sh}(Y)}(1) \to \mathrm{Sub}_{\mathbf{Sh}(X)}(1)$ – equivalently, by 1.3.15, to an order-preserving map $\mathcal{O}(Y) \to \mathcal{O}(X)$. Moreover, this map preserves finite intersections (since they are products in $\mathbf{Sh}(Y)$); and it has a right adjoint (since g_* also restricts to a functor between subobjects of 1) and hence preserves arbitrary unions. So it is a frame homomorphism; we let $F(g) \colon X \to Y$ denote the corresponding locale map. Now we noted earlier that the restriction of $G(f)^*$ to subterminal objects is just f^*, i.e. that $FG(f) = f$, for any locale map f. And for a geometric morphism g, we know that g^* is determined up to isomorphism by its restriction to subterminal objects, since every object of $\mathbf{LH}/Y$ can be written as a colimit of a diagram in $\mathrm{Sub}(1)$, and g^* preserves colimits; hence we have $GF(g) \cong g$. Similarly, if h is another geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, then a natural transformation $g^* \to h^*$ is uniquely determined by its values at

subterminal objects; hence there is at most one such transformation, and it exists iff $g^*(V) \leq h^*(V)$ for all $V \in \mathcal{O}(Y)$. So (F is functorial, and) the isomorphism $GF(g) \cong g$ is natural in g. □

From now on, we shall usually not distinguish notationally between a continuous map of locales and the geometric morphism of sheaf toposes which it induces; in other words, we shall identify **Loc** with the full sub-2-category of $\mathfrak{Top}$ whose objects are toposes of the form $\mathbf{Sh}(X)$. In fact we can do slightly more than this: since **Loc** has a terminal object 1, and it is easy to see (using any of the three descriptions of the last section) that $\mathbf{Sh}(1)$ is equivalent to **Set**, we can in fact regard $\mathbf{Loc} \cong \mathbf{Loc}/1$ as a full sub-2-category of $\mathfrak{Top}/\mathbf{Set}$.

The next problem is obviously to identify which objects of $\mathfrak{Top}/\mathbf{Set}$ lie in this sub-2-category; i.e. which toposes over **Set** are equivalent to ones of the form $\mathbf{Sh}(X)$. One advantage of working with locales rather than spaces is that the answer to this question takes on a particularly simple form. First we recall from A4.1.9 that, for a topos $\mathcal{E}$, there is (up to canonical isomorphism) at most one geometric morphism $\mathcal{E} \to \mathbf{Set}$, and that such a morphism exists iff $\mathcal{E}$ is locally small and has set-indexed copowers. The latter condition does not imply that $\mathcal{E}$ has all set-indexed coproducts, as we saw in Example A2.1.7, but it does imply the existence of some nontrivial coproducts:

Lemma 1.4.6 *If there exists a geometric morphism $\gamma\colon \mathcal{E} \to \mathbf{Set}$, then any set-indexed family of subterminal objects in $\mathcal{E}$ has a coproduct.*

Proof Let $(U_i \mid i \in I)$ be such a family. For each i, let $\chi_i\colon 1 \to \Omega$ be the classifying map of $U_i \rightarrowtail 1$; then since $\gamma^*(I)$ is the I-indexed copower of 1, the χ_i induce a map $\chi\colon \gamma^*(I) \to \Omega$. Let $m\colon U \rightarrowtail \gamma^*(I)$ be the subobject classified by χ. Then for each $i \in I$ we have a pullback

$$\begin{tikzcd} U_i \arrow[r, "\phi_i"] \arrow[d, tail] & U \arrow[d, tail, "m"] \\ 1 \arrow[r, "\nu_i"] & \gamma^*(I) \end{tikzcd}$$

where ν_i is the ith coprojection. But pullback along m has a right adjoint (A2.3.3) and so preserves any colimits which exist; so 1_U is the coproduct of the ϕ_i in $\mathcal{E}/U$, i.e. U is the coproduct of the U_i in $\mathcal{E}$. □

An easy generalization of the proof of 1.4.6 shows that, under the same hypothesis on $\mathcal{E}$, we may form the coproduct of an arbitrary set-indexed family $(A_i \mid i \in I)$ of objects of $\mathcal{E}$, provided there exists some $A \in \text{ob}\,\mathcal{E}$ which contains each A_i as a subobject.

We recall that in A4.6.1 we defined a geometric morphism $f: \mathcal{E} \to \mathcal{F}$ to be *localic* if every object of $\mathcal{E}$ is expressible as a quotient of a subobject of an object of the form f^*A, $A \in \text{ob } \mathcal{F}$. Now we can state

Theorem 1.4.7 *Let $\gamma: \mathcal{E} \to \mathbf{Set}$ be a topos defined over $\mathbf{Set}$. The following conditions are equivalent:*

(i) *$\mathcal{E}$ is equivalent to a topos of the form $\mathbf{Sh}(X)$, for some locale X.*

(ii) *The subterminal objects of $\mathcal{E}$ form a generating set.*

(iii) *The geometric morphism γ is localic.*

It was of course this theorem which inspired the use of the term 'localic' for geometric morphisms.

Proof (i) $\Rightarrow$ (ii): We already know that $\mathbf{Sh}(X)$ is generated by subterminal objects, since every object of $\mathbf{LH}/X$ is the colimit of a diagram whose vertices are subterminal objects. (Alternatively, we could argue that the representable functors form a generating set for $[\mathcal{O}(X)^{\text{op}}, \mathbf{Set}]$, and hence also for the full subcategory $\mathbf{Sh}(X)$ (since they lie in this subcategory); but they are precisely the subterminal objects of $\mathbf{Sh}(X)$, by 1.3.15.)

(ii) $\Rightarrow$ (iii): If the subterminal objects of $\mathcal{E}$ generate, then the family of all maps from such objects to a given object A of $\mathcal{E}$ is jointly epic; and this family is a set, since $\mathcal{E}$ is locally small. So by 1.4.6 we can construct an epimorphism $U \twoheadrightarrow A$ where U is a coproduct of subterminal objects, and hence (by the proof of 1.4.6) a subobject of some $\gamma^*(I)$. So γ is localic.

(iii) $\Rightarrow$ (ii): Conversely, if γ is localic, then for any object A we have an epimorphism $U \twoheadrightarrow A$ where U is a subobject of some $\gamma^*(I)$, and hence (by pulling back along the coprojections) an I-indexed coproduct of subterminal objects. So the family of all maps from subterminal objects to A is epic, i.e. the subterminal objects form a separating family for $\mathcal{E}$ – and in fact they form a set, since we can identify $\text{Sub}_{\mathcal{E}}(1)$ with $\gamma_*(\Omega_{\mathcal{E}})$. But $\mathcal{E}$ is balanced (A1.6.2), so any separating set is also a generating set.

(ii) $\Rightarrow$ (i): Suppose $\mathcal{E}$ is generated by subterminal objects. The lattice $\text{Sub}_{\mathcal{E}}(1)$ of subterminal objects of $\mathcal{E}$ is a Heyting algebra, by A1.5.10 (or by A1.6.3). But it also has set-indexed unions by 1.4.6, since we can construct these as images of maps from coproducts to 1; so by 1.1.2 it is a frame, and we take it to be $\mathcal{O}(X)$. Now for any object A of $\mathcal{E}$, the restriction of the functor $\mathcal{E}(-, A)$ to $\text{Sub}(1)$ is a presheaf $P(A)$ on X; and in fact it is a sheaf, since the construction of unions in $\mathcal{O}(X)$ just described enables us to show that they are colimits (in $\mathcal{E}$) of diagrams like that described in 1.3.4. Clearly, given a morphism $f: A \to B$ of $\mathcal{E}$, composition with f yields a natural transformation $P(A) \to P(B)$; so P is a functor $\mathcal{E} \to \mathbf{Sh}(X)$. In the opposite direction, given a sheaf F on X, we can regard the diagram $D(\Theta(F))$ constructed in the last section as a diagram in $\mathcal{E}$ (whose vertices are subterminal objects) rather than in $\mathbf{LH}/X$; and it has a colimit in $\mathcal{E}$ by 1.4.6, since we may construct colimits in the usual way from

coproducts and coequalizers. We take this colimit to be $Q(F)$; then the proof in 1.3.10 that Δ was functorial also shows that Q is a functor $\mathbf{Sh}(X) \to \mathcal{E}$.

Now if F is any sheaf on X, the proof given in 1.3.11 that $\Gamma\Delta\Theta(F) \cong F$ can be carried through (with suitable changes of notation) in $\mathcal{E}$ rather than in $\mathbf{LH}/X$, and it yields a proof that $PQ(F) \cong F$, naturally in F. So it remains to show that we have a natural isomorphism $QP(A) \cong A$, for $A \in \mathrm{ob}\ \mathcal{E}$. We certainly have a natural transformation $\epsilon\colon QP \to 1_{\mathcal{E}}$, since the family of all maps from subterminal objects to A forms a cone under the diagram whose colimit is $QP(A)$. And, by the construction of this diagram, each morphism from a subterminal object to A factors through ϵ_A. Moreover, the factorization is unique, since the argument used to prove $PQ(F) \cong F$ shows that any morphism $U \to QP(A)$, where U is subterminal, is one of the legs of the colimit cone. So the fact that the subterminal objects form a generating set implies that ϵ_A is an isomorphism for any A. □

Remark 1.4.8 If we consider how much of the proof of (ii) ⇒ (i) above survives without the assumption that $\mathcal{E}$ is generated by subterminal objects, we see that this assumption was only used in the proof that ϵ is an isomorphism. Hence, for a general **Set**-topos $\gamma\colon \mathcal{E} \to \mathbf{Set}$, the functor Q is left adjoint to P, and identifies $\mathbf{Sh}(X)$ (where $\mathcal{O}(X)$ is the frame of subterminal objects of $\mathcal{E}$) with a coreflective subcategory of $\mathcal{E}$. In fact it can be shown that Q preserves finite limits; so the pair (P, Q) is a geometric morphism, and this yields another construction of the hyperconnected–localic factorization of the geometric morphism γ, as originally considered in A4.6.5. In particular, we note once again that the sub-2-category $\mathfrak{LTop}/\mathbf{Set}$ of localic toposes over **Set** is reflective in $\mathfrak{Top}/\mathbf{Set}$ (cf. A4.6.12), and hence that *the functor* **Loc** $\to$ $\mathfrak{Top}/\mathbf{Set}$ *which sends* X *to* $\mathbf{Sh}(X)$ *preserves limits* (in an appropriate 2-categorical sense). This is one more advantage of working with locales rather than spaces: as we have already seen, the inclusion **Sob** $\to$ **Loc** fails to preserve even finite limits (the examples we saw involved intersections of subspaces, but there are also easy examples to show that a product of two spatial locales can fail to be spatial), and hence so does $X \mapsto \mathbf{Sh}(X)$ considered as a functor on **Sp**.

Remark 1.4.9 If we wish to characterize toposes of the form $\mathbf{Sh}(X)$ for a space (as opposed to a locale) X, we may do so by adding to the conditions of 1.4.7(ii) or (iii) the requirement that the family of all inverse image functors $\mathcal{E} \to \mathbf{Set}$ should be jointly conservative. We know that this is true for the topos of sheaves on a space, by the description of these functors in 1.4.2. Conversely, if Y is a non-spatial locale, then any point of Y factors through its spatial part Y_p, and so the corresponding inverse image functor $\mathbf{Sh}(Y) \to \mathbf{Set}$ inverts any inclusion between subterminal objects which is inverted by the frame homomorphism $\mathcal{O}(Y) \to \mathcal{O}(Y_p)$.

Before leaving this topic, we pause to note a very special case of 1.4.7, which fulfils a promise made at the end of Section A2.1.

Corollary 1.4.10 *For a* **Set**-*topos* $\gamma\colon \mathcal{E} \to \mathbf{Set}$, *the following are equivalent:*

(i) γ *is an equivalence.*

(ii) $\mathcal{E}$ *is non-degenerate (i.e.* $0 \to 1$ *is not an isomorphism), and its terminal object is a generator.*

Proof (i) $\Rightarrow$ (ii) since we know that **Set** has these properties.

(ii) $\Rightarrow$ (i): If (ii) holds, then by 1.4.7 $\mathcal{E}$ is equivalent to $\mathbf{Sh}(X)$, where X is the locale defined by $\mathcal{O}(X) = \mathrm{Sub}_{\mathcal{E}}(1)$. But if U is any subterminal object of $\mathcal{E}$ such that $U \rightarrowtail 1$ is not an isomorphism, then $0 \rightarrowtail U$ must be an isomorphism (since there are no morphisms from 1 to either 0 or U); so, given that $\mathcal{E}$ is non-degenerate, $\mathcal{O}(X)$ has just two elements, and X is the terminal object 1 of **Loc**. $\square$

C1.5 Some topological properties of toposes

In this section, our aim is to show how some familiar topological properties of locales and continuous maps are reflected in categorical properties of the corresponding sheaf toposes and geometric morphisms. This subject will be investigated in much greater detail in Chapter C3 below; but it will be useful to introduce some of the definitions now, before we move on to considering sheaves on sites. In fact, despite the title of the section, almost all the properties we study will be properties of geometric morphisms rather than properties of toposes: it turns out that many things which are traditionally thought of as properties of a space (or locale) X are actually properties of the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Set}$. Thus the act of translating these properties into their topos-theoretic form allows us to 'relativize' them; that is, to apply them to arbitrary geometric morphisms and thence, by specializing to the case of morphisms $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$, to arbitrary continuous maps of locales. We shall investigate this point of view further in Section C1.6.

First, and most importantly, we show how sheaf toposes motivate the definitions of geometric inclusions and surjections, introduced in Section A4.2. One half of the following result was in fact proved there (see A4.2.12(c)), but the other was deliberately postponed until now.

Proposition 1.5.1 *Let* $f\colon X \to Y$ *be a continuous map of locales.*

(i) f *is a sublocale inclusion (i.e. a regular monomorphism in* **Loc**) *iff the direct image functor* $f_*\colon \mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ *is full and faithful.*

(ii) f *is an epimorphism in* **Loc** *iff the inverse image functor* $f^*\colon \mathbf{Sh}(Y) \to \mathbf{Sh}(X$ *is faithful.*

Proof (i) f is a sublocale inclusion iff the frame homomorphism $f^*\colon \mathcal{O}(Y) \to \mathcal{O}(X)$ is surjective. If this condition holds, then it is easy to see that, for functors $F, G\colon \mathcal{O}(Y)^{\mathrm{op}} \rightrightarrows \mathbf{Set}$, any natural transformation $F \circ f^* \to G \circ f^*$

derives from a unique natural transformation $F \to G$; so f_* is full and faithful. Conversely, if f_* is full and faithful, then the counit of $(f^* \dashv f_*)$ is an isomorphism; so any subterminal object U of $\mathbf{Sh}(X)$ satisfies $U \cong f^*f_*(U)$, i.e. the restriction of f^* to subterminal objects is surjective.

(ii) This requires a little care, because in general epimorphisms are not stable under pullback in **Loc**, as we saw in 1.2.12. However, it follows from 1.2.11 that epimorphisms in **Loc** are stable under pullback along open inclusions. From here, a standard argument along the lines of 1.3.2(ii) shows that they are stable under pullback along local homeomorphisms. Now if

$$\begin{tikzcd} E \arrow[rr, shift left, "h"] \arrow[rr, shift right, "k"'] \arrow[dr, "p"'] & & F \arrow[dl, "q"] \\ & Y & \end{tikzcd}$$

is a parallel pair of morphisms in $\mathbf{LH}/Y$ having the same image under f^*, then h and k are equalized by the projection $f^*(E) \twoheadrightarrow E$, and so if f is epimorphic they must be equal.

Conversely, if $f^* : \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ is faithful then it is also conservative, since $\mathbf{Sh}(Y)$ is balanced; so in particular it reflects isomorphisms between subterminal objects, and hence $f^* : \mathcal{O}(Y) \to \mathcal{O}(X)$ is injective, i.e. f is an epimorphism in **Loc**. □

Corollary 1.5.2 *Let $f : X \to Y$ be a continuous map of locales. Then the image factorization of the induced geometric morphism* $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$*, as constructed in* A4.2.10*, corresponds to the factorization of f in* **Loc** *as an epimorphism followed by a regular monomorphism.* □

For spaces, the analogue of 1.5.1 does not look quite so neat as it does for locales. If $f : X \to Y$ is a subspace inclusion (resp. a continuous surjection), then the corresponding map of locales is a regular monomorphism (resp. an epimorphism); but the converses do not hold without mild separation hypotheses on the spaces involved. For the inclusion case, all we require is that X should be a T_0-space, which is covered by our blanket assumption of sobriety (if $f^{-1} : \mathcal{O}(Y) \to \mathcal{O}(X)$ is surjective, then two points of X with the same image under f must lie in exactly the same open sets of X); but in the surjection case, the most that we can deduce from knowing that f^{-1} is injective is that the image of f is *superdense* in Y, i.e. that its subclosure (as defined in 1.2.5) is the whole of Y. If Y is a T_D-space, this forces f to be surjective; but in general it does not. Nevertheless, we shall commonly use the term *surjection* to refer to an epimorphism in **Loc**.

Next, we consider open maps. Of course, we say a continuous map $f : X \to Y$ of locales is *open* if, for every open sublocale U of X, its image $f_!(U)$ in $\mathrm{Sub}(Y)$

is an open sublocale of Y. Once again, this is not quite equivalent to the usual notion of openness for continuous maps of spaces; it is implied by the latter, but in the opposite direction it only implies that the set-theoretic image of each open subset of X is 'almost open' in Y, in the sense that its subclosure is open. Even when X and Y are both sober, this does not suffice to make all such images open (see [292]), but it obviously does so if Y is a T_D-space.

We shall need the following result from locale theory.

Lemma 1.5.3 *For a continuous map $f: X \to Y$ of locales, the following are equivalent:*

(i) *f is an open map.*

(ii) *The frame homomorphism $f^*: \mathcal{O}(Y) \to \mathcal{O}(X)$ has a left adjoint $f_!$, which satisfies the* Frobenius reciprocity *condition that*

$$f_!(U \cap f^*(V)) = f_!(U) \cap V$$

for all $U \in \mathcal{O}(X)$, $V \in \mathcal{O}(Y)$.

(iii) *$f^*: \mathcal{O}(Y) \to \mathcal{O}(X)$ is a homomorphism of complete Heyting algebras; i.e. it preserves arbitrary meets and the Heyting implication operation.*

Proof (i)$\Leftrightarrow$(ii): The assertion that f^* has a left adjoint is equivalent to saying that, for every $U \in \mathcal{O}(X)$, there is a unique smallest open sublocale $V = f_\bullet(U)$ of Y such that $U \rightarrowtail X \to Y$ factors through $V \rightarrowtail Y$. For openness, we need to know in addition that $f_\bullet(U)$ actually coincides with $f_!(U)$; equivalently, that the induced map $\check{f}: U \to f_\bullet(U)$ is a locale epimorphism. But if we identify the frames $\mathcal{O}(U)$ and $\mathcal{O}(f_\bullet(U))$ with the appropriate principal ideals in $\mathcal{O}(X)$ and $\mathcal{O}(Y)$, then the frame homomorphism $\check{f}^*$ sends an open $V \leq f_\bullet(U)$ to $f^*(V) \cap U$, and this is easily seen to have a left adjoint, namely $f_\bullet$ restricted to open sublocales contained in U. So the Frobenius reciprocity condition implies that the counit of this adjunction is the identity, and hence that $\check{f}^*$ is injective. Conversely, if $\check{f}^*$ is injective, then the Frobenius reciprocity condition holds for pairs (U, V) such that $V \leq f_\bullet(U)$; to deduce the general case, we note that $U \leq f^* f_\bullet(U)$ by adjointness, and so

$$f_\bullet(U \cap f^*(V)) = f_\bullet(U \cap f^* f_\bullet(U) \cap f^*(V)) = f_\bullet(U \cap f^*(f_\bullet(U) \cap V)) = f_\bullet(U) \cap V,$$

the last step by the special case already established.

(ii)$\Leftrightarrow$(iii): f^* preserves arbitrary meets iff it has a left adjoint, by the adjoint functor theorem; and then it also preserves implication iff the Frobenius reciprocity condition holds, by A1.5.8. $\square$

Theorem 1.5.4 *Let $f: X \to Y$ be a continuous map of locales. Then f is open iff the inverse image functor $f^*: \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ is a Heyting functor (i.e. commutes with universal quantification).*

Proof Suppose f is open. Then the same is true of the pullback of f along any open inclusion $V \rightarrowtail Y$ (since, by 1.2.10, this is the restriction of f to an open sublocale of its domain), and hence it is also true for the pullback of f along any local homeomorphism $E \to Y$: if U is an open sublocale of $f^*(E)$, we can express it as a union of open sublocales $U \cap f^*(V)$ where $V \rightarrowtail E$ is an open sublocale such that $V \rightarrowtail E \to Y$ is an open inclusion, and the image of U under the morphism $f^*(E) \to E$ is the union of the images of the $U \cap f^*(V)$, and so open in E. (In fact we shall see later that open maps of locales are stable under arbitrary pullback.) Thus, for any object A of $\mathbf{Sh}(Y)$, we may deduce from 1.5.3(ii) that the functor $\mathrm{Sub}_{\mathbf{Sh}(Y)}(A) \to \mathrm{Sub}_{\mathbf{Sh}(X)}(f^*(A))$ obtained by applying f^* to subobjects of A has a left adjoint $f_!^A$. Moreover, the Frobenius reciprocity condition is tantamount to saying that these left adjoints commute with pullback along open inclusions (that is, monomorphisms in $\mathbf{Sh}(Y)$); and a similar argument to the one above enables us to show that they commute with pullback along arbitrary local homeomorphisms, i.e. the square

$$\begin{tikzcd}
\mathrm{Sub}(f^*(B)) \arrow[r, "f_!^B"] \arrow[d, "(f^*h)^*"] & \mathrm{Sub}(B) \arrow[d, "h^*"] \\
\mathrm{Sub}(f^*(A)) \arrow[r, "f_!^A"] & \mathrm{Sub}(A)
\end{tikzcd}$$

commutes up to isomorphism for any $h \colon A \to B$ in $\mathbf{Sh}(Y)$. Taking right adjoints of all the functors in this diagram, we deduce that

$$\begin{tikzcd}
\mathrm{Sub}(A) \arrow[r, "f^*"] \arrow[d, "\forall_h"] & \mathrm{Sub}(f^*(A)) \arrow[d, "\forall_{f^*h}"] \\
\mathrm{Sub}(B) \arrow[r, "f^*"] & \mathrm{Sub}(f^*(B))
\end{tikzcd}$$

also commutes up to isomorphism, i.e. f^* is a Heyting functor.

Conversely, suppose f^* is a Heyting functor. Since the Heyting implication in the subobject lattices of a Heyting category is definable in terms of universal quantification (A1.4.13), f^* also preserves this operation; in particular, $f^* \colon \mathcal{O}(Y) \to \mathcal{O}(X)$ is a Heyting algebra homomorphism. But if $\gamma \colon \mathbf{Sh}(Y) \to \mathbf{Set}$ is the unique geometric morphism, then $\gamma^*(I)$ is the I-indexed copower of 1 in $\mathbf{Sh}(Y)$, and its subobjects correspond to I-indexed families of subobjects of 1; and universal quantification along $\gamma^*(I) \to 1$ corresponds to taking intersections of such families. So f^* preserves all such intersections, and we have verified condition (iii) of 1.5.3. $\square$

In view of this result, we define an arbitrary geometric morphism $f\colon \mathcal{E} \to \mathcal{F}$ to be *open* if its inverse image is a Heyting functor. Open geometric morphisms will be studied in detail in Section C3.1.

Next, we consider compactness. Of course, a locale X is said to be *compact* if the top element X of $\mathcal{O}(X)$ is inaccessible by directed joins: i.e. if, whenever we have a directed family $(U_i \mid i \in I)$ of open sublocales of X whose union is the whole of X, then some U_i is the whole of X. This translates easily into topos-theoretic terms:

Lemma 1.5.5 *A locale X is compact iff the direct image functor $\gamma_*\colon \mathbf{Sh}(X) \to \mathbf{Set}$ preserves directed unions of subobjects.*

Proof We recall that γ_* is (isomorphic to) the representable functor $\mathbf{Sh}(X)(1, -)$. Now, given an object A of $\mathbf{Sh}(X)$ which is the directed union of a family of subobjects $(A_i \mid i \in I)$, for each $(x\colon 1 \to A) \in \gamma_*(A)$ we may form the pullback

$$\begin{tikzcd} U_i \arrow[r, tail] \arrow[d] & 1 \arrow[d, "x"] \\ A_i \arrow[r, tail] & A \end{tikzcd} ,$$

and then the U_i form a directed family of subterminal objects whose union is 1. So if X is compact, one of the U_i must be the whole of 1; i.e. x must factor through some $A_i \rightarrowtail A$. So $\gamma_*(A)$ is the union of the $\gamma_*(A_i)$. The converse follows immediately from considering subobjects of 1. □

We shall study morphisms whose direct image functors preserve directed unions of subterminal objects in Section C3.2, under the name *proper maps*; we shall see that there is an important sense in which the class of proper maps is 'dual' to that of open maps.

For not-necessarily-localic morphisms, the condition that a direct image functor should preserve directed unions seems less 'natural' than the condition that it should preserve all filtered colimits; however, the latter is strictly stronger (even in the localic case). We shall study geometric morphisms with this property, under the name *tidy morphisms*, in Section C3.4.

We digress briefly to consider a still stronger form of compactness. By a *focal point* of a space X, we mean a point x whose only neighbourhood is the whole of X. (The name is justified by the fact that a point is focal iff every sequence converges to it.) Clearly, in a T_0-space (in particular, in a sober space) there can be at most one focal point. We say a space is *local* if it has a focal point; this name derives from the fact that the prime spectrum of a local ring is a local space (the focal point being that which corresponds to the unique maximal ideal of the ring). An equivalent frame-theoretic condition is that the top element of $\mathcal{O}(X)$ is inaccessible by arbitrary joins (i.e. that $\{X\}$ is a completely prime filter

in $\mathcal{O}(X)$ – equivalently, that every open cover of X has a singleton subcover), and we take this as the definition of localness for a locale.

Lemma 1.5.6 *A locale X is local iff the direct image functor $\gamma_* \colon \mathbf{Sh}(X) \to \mathbf{Set}$ has a right adjoint.*

Proof If γ_* has a right adjoint, then it must be the *inverse* image part of a geometric morphism $c \colon \mathbf{Set} \to \mathbf{Sh}(X)$, and the latter must correspond by 1.4.5 to a point of the locale X. The completely prime filter in $\mathcal{O}(X)$ corresponding to this point is the set $\{U \in \mathcal{O}(X) \mid \gamma_*(U) = 1\}$; but from the description of γ_* as a representable functor, this set is the singleton $\{X\}$. So if such a point exists then X is local. Conversely, if X is local, then from the description of stalk functors given in 1.4.2 we see that the stalk functor at the focal point is isomorphic to $\mathbf{Sh}(X)(1, -)$; so the latter has a right adjoint. □

We call a geometric morphism $f \colon \mathcal{E} \to \mathcal{F}$ *local* if f_* has a right adjoint $f^\#$ which is full and faithful. (Note that the latter condition is automatic in the situation considered in 1.5.6, since the point c must be an inclusion.) Local morphisms will be considered in Section C3.6.

Next, we consider global and local connectedness. Traditionally, a space X is called *connected* if $\mathcal{O}(X)$ has precisely two complemented elements $\emptyset$ and X. As our definition for locales, we shall take the classically equivalent, but constructively more sensible, condition that, if $(U_i \mid i \in I)$ is a pairwise-disjoint family of open sublocales whose union is the whole of X, then exactly one U_i must be the whole of X. (For the equivalence, note that for each $i \in I$ the open U_i is complemented by the union of the rest of the U_j; so it must be either $\emptyset$ or X.)

Lemma 1.5.7 *For a locale X, the following are equivalent:*

(i) *X is connected.*

(ii) *The terminal object of $\mathbf{Sh}(X)$ is indecomposable (cf.* A1.1.10*).*

(iii) *The inverse image functor $\gamma^* \colon \mathbf{Set} \to \mathbf{Sh}(X)$ is full and faithful.*

(iv) *The direct image functor $\gamma_* \colon \mathbf{Sh}(X) \to \mathbf{Set}$ preserves coproducts.*

Proof (i) $\Leftrightarrow$ (ii) is trivial, given the identification of subterminal objects of $\mathbf{Sh}(X)$ with elements of $\mathcal{O}(X)$ (1.3.15).

(ii) $\Rightarrow$ (iv): Since γ_* is the functor represented by the terminal object of $\mathbf{Sh}(X)$, an element of $\gamma_*(\coprod_{i \in I} A_i)$ corresponds to a morphism $1 \to \coprod_{i \in I} A_i$, and hence (by pulling back the coprojections) to an I-indexed coproduct decomposition of 1 in $\mathbf{Sh}(X)$. So (ii) says that every element of $\gamma_*(\coprod_{i \in I} A_i)$ derives from an element of $\gamma_*(A_i)$ for some i; i.e. the former is the disjoint union of the latter.

(iv) $\Rightarrow$ (iii): If γ_* preserves coproducts, then so does the composite $\gamma_*\gamma^*$. But this composite also preserves 1; and every set I is an I-fold coproduct of copies of 1, so we have $\gamma_*\gamma^*(I) \cong \coprod_{i \in I} 1 \cong I$. Thus the counit of the adjunction $(\gamma^* \dashv \gamma_*)$ is an isomorphism; equivalently, γ^* is full and faithful.

(iii) $\Rightarrow$ (ii): Suppose given a coproduct decomposition $1 = \coprod_{i \in I} U_i$ in $\mathbf{Sh}(X)$. Then the inclusions $U_i \to 1$ combine to form a morphism $1 \cong \gamma^*(1) \to \gamma^*(I)$. If this equals $\gamma^*(i)$ for some $i: 1 \to I$ in **Set**, then the corresponding U_i is the whole of X. □

We therefore define a geometric morphism to be *connected* if its inverse image part is full and faithful, or equivalently if the unit of the adjunction is an isomorphism. We note that hyperconnected morphisms, as defined in A4.6.6, are connected; and so are local morphisms as defined earlier in this section, since $f_* f^*$ is isomorphic to the identity functor iff its right adjoint $f_* f^{\#}$ is. In 3.4.14 below we shall prove that 1.5.7(iii) and (iv) are equivalent over an arbitrary base, provided we interpret the coproducts in (iv) as being indexed by objects of the base.

We say a locale X is *locally connected* if every open sublocale of X is expressible as a union of connected open sublocales. It is easy to see that, when restricted to spatial locales, this is equivalent to any of the usual definitions of local connectedness for spaces. Also, it is a local property: that is, not only is it inherited by open sublocales, but if X has an open cover by locally connected sublocales V_i, then X is locally connected (since any $U \in \mathcal{O}(X)$ is the union of the sublocales $U \cap V_i$, and these are unions of connected open sublocales). Hence if X is locally connected, so is the domain of any local homeomorphism $p: E \to X$. We shall also require

Lemma 1.5.8 *If X is locally connected, then any open sublocale of X has a pairwise-disjoint covering by connected open sublocales.*

Proof It suffices to construct such a covering of X itself. Let $(U_i \mid i \in I)$ be an arbitrary family of connected open sublocales covering X; we define an equivalence relation R on I to be the transitive closure of the relation which relates i and i' precisely when $U_i \cap U_{i'} \neq \emptyset$. Let J be the set of R-equivalence classes, and for each $j \in J$ let $V_j = \bigcup_{i \in j} U_i$; then by construction (plus the infinite distributive law in $\mathcal{O}(X)$) the V_j are pairwise disjoint, and they cover X. We must show that they are connected. But if $(W_k \mid k \in K)$ is a pairwise-disjoint open cover of some V_j, then for each $i \in j$ the sublocales $(W_k \cap U_i \mid k \in K)$ form a pairwise-disjoint open cover of U_i, and so one of them, say $W_{f(i)} \cap U_i$, must be the whole of U_i. This defines a function $f: j \to K$; but if $U_i \cap U_{i'} \neq \emptyset$ we must have $f(i) = f(i')$, since the W_k are disjoint, and so f is constant on the equivalence class j. Hence there exists $k \in K$ such that W_k contains every U_i, $i \in j$, and so $W_k = V_j$. □

It is easy to see that the disjoint decomposition of X into connected open sublocales, which we have constructed in the proof of 1.5.8, is unique; of course, we call its members the *components* of X. Although the proof of 1.5.8 appears to require classical logic, in that the equivalence relation R is defined in terms of an inequality, we shall see when we reach Section C3.3 that the definition of

local connectedness implies the (classically vacuous) property of openness (i.e. the property that the unique map from X to the terminal locale 1 is open), and for open locales we may use the notion of positivity developed in 3.1.17 as a substitute for nonemptiness. Thus the constructively valid version of 1.5.8 asserts that, for any locally connected locale X, there is a family of connected open sublocales $\{U_i \mid i \in I\}$ which is 'pairwise disjoint' in the sense that $U_i \cap U_j$ is positive iff $i = j$, and whose union is the whole of X.

Proposition 1.5.9 *For a locale X, the following are equivalent:*

(i) *X is locally connected.*

(ii) *The geometric morphism $\gamma\colon \mathbf{Sh}(X) \to \mathbf{Set}$ is essential, i.e. γ^* has a left adjoint $\gamma_!$.*

(iii) *$\gamma^* : \mathbf{Set} \to \mathbf{Sh}(X)$ is a cartesian closed functor (i.e. preserves exponentials).*

Proof (i) $\Rightarrow$ (ii): We define $\gamma_!$ in terms of local homeomorphisms; specifically, we define $\gamma_!(p\colon E \to X)$ to be the set of components of E (which we have already observed to be locally connected). This is functorial, because the image of a connected locale under a continuous map is readily seen to be connected, and so a map $f\colon E \to F$ between (the domains of) two local homeomorphisms over X maps each component of E into a component of F. Now suppose given a morphism $f\colon (p\colon E \to X) \to \gamma^*(I)$ in $\mathbf{LH}/X$. Since $\gamma^*(I)$ is the disjoint union of the subobjects $\nu_i : 1 \rightarrowtail \gamma^*(I)$, the inverse images of these subobjects along f form a disjoint decomposition of E, and so each component of E is contained in just one of them; this defines a function $\gamma_!(p) \to I$. Conversely, given such a function g, we may define a morphism $p \to \gamma^*(I)$ in $\mathbf{LH}/X$, by setting its restriction to the kth component E_k of E equal to the composite

$$E_k \longrightarrow 1 \xrightarrow{\nu_{g(k)}} \gamma^*(I)\ .$$

These two constructions are clearly inverse to each other, and define the required adjunction $(\gamma_! \dashv \gamma^*)$.

(ii) $\Rightarrow$ (iii): Given that γ^* has a left adjoint, we know from A1.5.8 that it is cartesian closed iff the Frobenius reciprocity condition

$$\gamma_!(E \times \gamma^*(I)) \cong \gamma_!(E) \times I$$

holds for all objects E of $\mathbf{Sh}(X)$, I of $\mathbf{Set}$. But $E \times \gamma^*(I)$ is the coproduct of I copies of E in $\mathbf{Sh}(X)$, since $E \times (-)$ preserves coproducts; and $\gamma_!$ also preserves coproducts since it is a left adjoint, so this is immediate.

(iii) $\Rightarrow$ (i): For any object E of $\mathbf{Sh}(X)$, morphisms $E \to \gamma^*(I)$ correspond to I-indexed coproduct (= disjoint union) decompositions of E in $\mathbf{Sh}(X)$. So the assertion that γ^* preserves the exponential J^I is equivalent to saying that, for any E, the map from the set of J^I-indexed decompositions of E to the set of I-indexed

families of J-indexed decompositions of E, which sends a decomposition $(E_f \mid f \in J^I)$ to the family of decompositions in which the jth member of the ith family is $\coprod_{f(i)=j} E_f$, is bijective. But if this holds for all I and J, then the lattice $\mathrm{Sub}^c(E)$ of complemented subobjects of E is complete, since given any family of complemented subobjects $(E_i \mid i \in I)$ we can form the I-indexed family of $\{0,1\}$-indexed decompositions consisting of the E_i and their complements, and then one of the elements of the $\{0,1\}^I$-indexed decomposition corresponding to this must be the meet of the E_i in $\mathrm{Sub}^c(E)$. And a similar argument shows that $\mathrm{Sub}^c(E)$ must satisfy the complete distributive law; but since it is a Boolean algebra, this forces it to be atomic, i.e. every object of $\mathbf{Sh}(X)$ has a coproduct decomposition into subobjects which admit no nontrivial decompositions. In particular, this holds for every subterminal object in $\mathbf{Sh}(X)$; but this is precisely the statement that X is locally connected. □

Proposition 1.5.9 illustrates one of the problems we encounter in seeking to 'relativize' topological properties: it is sometimes the case that there are several conditions which are equivalent for morphisms of the form $\mathbf{Sh}(X) \to \mathbf{Set}$, but not for more general geometric morphisms. In this case, the conditions in (ii) and (iii) are not equivalent for a general $f : \mathcal{F} \to \mathcal{E}$: the mere existence of a left adjoint for f^* does not suffice to ensure that Frobenius reciprocity holds, and the mere preservation of exponentials by f^* does not suffice to ensure that it has a left adjoint. The answer in this case (as in many others) is to demand that the appropriate condition should hold in the context of $\mathcal{E}$-indexed categories (where $\mathcal{E}$ is the codomain of f): thus we define a geometric morphism f to be *locally connected* if f^* has an $\mathcal{E}$-indexed left adjoint (and this *is* equivalent to saying that it is cartesian closed as an $\mathcal{E}$-indexed functor, as we shall see in 3.3.1 below).

Before concluding this section, we feel bound to give one example of a topological property which is 'absolute' in the sense that it really is a property of the topos $\mathbf{Sh}(X)$ rather than of the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Set}$. The best-known such property is extremal disconnectedness: recall that a topological space X is said to be *extremally disconnected* if the closure of each open subset of X is open. An equivalent formulation, which makes sense for locales as well as spaces, is that the Heyting negation map $\neg : \mathcal{O}(X) \to \mathcal{O}(X)$ takes values in the sublattice of complemented elements of $\mathcal{O}(X)$; equivalently again, the fixset $\mathcal{O}(X)_{\neg\neg}$ corresponding to the smallest dense sublocale of X (cf. 1.2.6(c)) coincides with the sublattice of complemented elements of $\mathcal{O}(X)$.

Lemma 1.5.10 *A locale X is extremally disconnected iff the canonical morphism $(\top, \bot) : 1 \amalg 1 \to \Omega_{\neg\neg}$ is an isomorphism in $\mathbf{Sh}(X)$.*

Proof The object $1 \amalg 1$ of $\mathbf{Sh}(X)$ is the sheaf of sections of the codiagonal map $X \amalg X \to X$; so its sections over an open sublocale U correspond to complemented elements of $\mathcal{O}(U)$. Similarly, since elements of $\Omega(U)$ are arbitrary open sublocales of U, the elements of $\Omega_{\neg\neg}(U)$ are the regular open sublocales of U, i.e. those which coincide with the interior of their closure in U. So the assertion that

$1 \amalg 1 \to \Omega_{\neg\neg}$ is an isomorphism is equivalent to saying that every open sublocale of X is extremally disconnected. But it is easy to verify that extremal disconnectedness is inherited by open sublocales; so this happens iff X is extremally disconnected. □

Toposes satisfying the condition in Lemma 1.5.10 are studied in greater detail in Section D4.6.

Suggestions for further reading: Grayson [417, 418], Johnstone [520, 523, 524], Joyal & Tierney [560].

C1.6 Internal locales

The concept of frame, as defined in Section C1.1, is clearly capable of being interpreted constructively; and so we can talk about internal frames and locales in any topos, in particular in a topos $\mathbf{Sh}(X)$ of sheaves on a locale. The main result to be proved in this section is that the category $\mathbf{Loc}(\mathbf{Sh}(X))$ of internal locales in $\mathbf{Sh}(X)$ is equivalent to the slice category $\mathbf{Loc}/X$ of locales over X; so any constructively valid result about locales, when interpreted in $\mathbf{Sh}(X)$, yields a result about locales over X. (This is, once again, an instance of the superiority of locales over spaces: we can also talk about internal topological spaces in $\mathbf{Sh}(X)$, but these correspond to a rather restricted class of spaces over X, as we shall see in 1.6.7 below.) For this section, we shall assume some familiarity with the interpretation of logic in a topos, as set out in Chapter D1, and with the notion of completeness for internal posets (cf. B2.3.9).

We recall first from A2.1.8 that the subobject classifier Ω_Y in $\mathbf{Sh}(Y)$ is the sheaf given by $\Omega_Y(U) = \{V \in \mathcal{O}(Y) \mid V \leq U\}$ for each $U \in \mathcal{O}(Y)$. As it is in any topos (A1.6.3), Ω is an internal Heyting algebra in $\mathbf{Sh}(Y)$; it is also complete as an internal poset in $\mathbf{Sh}(Y)$, by B2.3.8(a). Now suppose we are given a continuous map $f\colon Y \to X$ of locales, inducing a geometric morphism $f\colon \mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ as in 1.4.4; then the direct image functor f_* preserves Heyting algebras (because it is cartesian) and complete internal posets (by B2.3.7), and so the sheaf $f_*(\Omega_Y)$ on X (defined by $f_*(\Omega_Y)(U) = \{V \in \mathcal{O}(Y) \mid V \leq f^*(U)\}$) is an internal complete Heyting algebra, and hence an internal frame, in $\mathbf{Sh}(X)$. Let us write $\mathcal{I}(f)$ for the corresponding internal locale in $\mathbf{Sh}(X)$, and call it the *internalization* of f.

Proposition 1.6.1 *The assignment $f \mapsto \mathcal{I}(f)$ is a functor* $\mathbf{Loc}/X \to \mathbf{Loc}(\mathbf{Sh}(X))$.

Proof Suppose given a commutative triangle of locale maps

$$\begin{tikzcd} Y \arrow[rr, "h"] \arrow[dr, "f"'] & & Z \arrow[dl, "g"] \\ & X & \end{tikzcd}$$

Then for each $U \in \mathcal{O}(X)$, h^* restricts to a frame homomorphism $h_U^* : g_*(\Omega_Z)(U) \to f_*(\Omega_Y)(U)$, since $h^*g^*(U) = f^*(U)$. Moreover, these maps form a natural transformation $g_*(\Omega_Z) \to f_*(\Omega_Y)$, since h^* commutes with binary intersections. This natural transformation is clearly an internal lattice homomorphism in $\mathbf{Sh}(X)$; to prove that it is an internal frame homomorphism, it is enough to show that the right adjoints of the h_U^* also form a natural transformation. But the right adjoint of h_U^* is easily seen to send $V \leq f^*(U)$ to $h_*(V) \cap g^*(U)$, and this is natural in U since

$$\begin{aligned} h_*(V \cap f^*(U')) \cap g^*(U') &= h_*(V) \cap h_*h^*g^*(U') \cap g^*(U') \\ &= h_*(V) \cap g^*(U') \\ &= (h_*(V) \cap g^*(U)) \cap g^*(U') \end{aligned}$$

whenever $U' \leq U$ in $\mathcal{O}(X)$. Thus we have an internal locale morphism $\mathcal{I}(h) : \mathcal{I}(f) \to \mathcal{I}(g)$; and it is clearly functorial in h. □

In the opposite direction, suppose we are given an internal frame L in $\mathbf{Sh}(X)$. Since the global sections functor $F \mapsto F(X)$ is the direct image of the (unique) geometric morphism $\mathbf{Sh}(X) \to \mathbf{Set}$ (cf. A4.1.9), it follows from an argument we have already used in this section that $L(X)$ is a complete Heyting algebra, and hence a frame, in $\mathbf{Set}$. Moreover, since we may identify elements of $\mathcal{O}(X)$ with subterminal objects in $\mathbf{Sh}(X)$, and since elements of $L(U)$ correspond to morphisms $U \to L$ in $\mathbf{Sh}(X)$, the completeness of L as an internal poset tells us that, for each $V \leq U$ in $\mathcal{O}(X)$, the restriction map $\rho_V^U : L(U) \to L(V)$ has a left adjoint $\sigma_U^V : L(V) \to L(U)$, and that the (Beck–Chevalley) square

$$\begin{tikzcd} L(V_1) \arrow[r, "\rho"] \arrow[d, "\sigma"] & L(V_1 \cap V_2) \arrow[d, "\sigma"] \\ L(U) \arrow[r, "\rho"] & L(V_2) \end{tikzcd}$$

commutes whenever $V_1, V_2 \leq U$ (cf. B1.4.2). (In particular, taking $V_1 = V_2$, this says that $\rho_V^U \sigma_U^V$ is the identity on $L(V)$.) Since L is an internal Heyting algebra

in $\mathbf{Sh}(X)$, the maps ρ^U_V are also Heyting algebra homomorphisms, from which we deduce by A1.5.8 that the Frobenius reciprocity law

$$\sigma^V_U(y \wedge \rho^U_V(x)) = \sigma^V_U(y) \wedge x$$

holds for $x \in L(U)$, $y \in L(V)$. In particular, taking $U = X$ and y to be the top element $\top_V$ of $L(V)$, we obtain

$$\sigma^V_X \rho^X_V(x) = x \wedge s(V)$$

where $s(V) = \sigma^V_X(\top_V) \in L(X)$. It follows easily that $L(V)$ may be identified, via σ^V_X, with the principal ideal $\{y \in L(X) \mid y \leq s(V)\}$ of $L(X)$.

We claim that $s\colon \mathcal{O}(X) \to L(X)$ is a frame homomorphism. To see this, note first that if we chase the element $\top_{V_1}$ around the Beck–Chevalley square above (with $U = X$) and then apply $\sigma^{V_2}_X$ to the result, we obtain the equality

$$s(V_1 \cap V_2) = \sigma^{V_2}_X \rho^X_{V_2}(s(V_1)) = s(V_1) \wedge s(V_2),$$

so s preserves binary meets (and clearly $s(X) = \top_X$). Now suppose $U = \bigcup_{i \in I} U_i$ in $\mathcal{O}(X)$. Then $\top_U = \bigvee_{i \in I} \sigma^{U_i}_U(\top_{U_i})$ in $L(U)$, since L is a sheaf and each side of the equation restricts to $\top_{U_i}$ in $L(U_i)$, for each i. Applying σ^U_X to this equation (and noting that σ^U_X, being a left adjoint, preserves joins), we obtain $s(U) = \bigvee_{i \in I} s(U_i)$, so s preserves joins.

Thus we have proved

Proposition 1.6.2 *For any internal frame L in $\mathbf{Sh}(X)$, there exists a locale morphism $f\colon Y \to X$ such that $L \cong \mathcal{O}(\mathcal{I}(f))$.*

Proof Take Y to be the locale defined by setting $\mathcal{O}(Y) = L(X)$, and f to be the locale map defined by setting $f^* = s$. Then for each $U \in \mathcal{O}(X)$ we have $L(U) \cong \{x \in L(X) \mid x \leq f^*(U)\}$, as already noted; it is easy to check that these isomorphisms are natural in U, and so define the required isomorphism of internal frames in $\mathbf{Sh}(X)$. $\square$

To complete the proof of the equivalence, we must make the assignment $L \mapsto (f\colon Y \to X)$ of Proposition 1.6.2 into a functor $\mathcal{E}\colon \mathbf{Loc}(\mathbf{Sh}(X)) \to \mathbf{Loc}/X$ ($\mathcal{E}$ stands for 'externalization', of course). But this is straightforward: an internal frame homomorphism $h^*\colon L \to L'$ in $\mathbf{Sh}(X)$ will induce an external frame homomorphism $h^*(X)\colon L(X) \to L'(X)$; and because the diagram

$$\begin{tikzcd}
L(U) \arrow[r, "h^*(U)"] \arrow[d, "\sigma^U_X"] & L'(U) \arrow[d, "\sigma^U_X"] \\
L(X) \arrow[r, "h^*(X)"] & L'(X)
\end{tikzcd}$$

commutes (since h^* preserves joins) and $h^*(U)$ preserves the top element, we have $h^*(X)(s(U)) = s'(U)$, i.e. $h^*(X)$ is a frame homomorphism under $\mathcal{O}(X)$. Once again, it is trivial to verify that this construction is functorial; so we have

Theorem 1.6.3 *For any locale X, the categories* $\mathbf{Loc}(\mathbf{Sh}(X))$ *and* $\mathbf{Loc}/X$ *are equivalent.*

Proof We have constructed the functors $\mathcal{I}$ and $\mathcal{E}$, and we verified in 1.6.2 that $\mathcal{I}\mathcal{E}$ is isomorphic to the identity on $\mathbf{Loc}(\mathbf{Sh}(X))$ (we did not verify that the isomorphism is natural, but that is entirely straightforward); so it remains to show that $\mathcal{E}\mathcal{I}$ is naturally isomorphic to the identity on $\mathbf{Loc}/X$. It is clear from the constructions that for any $f: Y \to X$ we have $\mathcal{O}(\mathrm{dom}\,\mathcal{E}\mathcal{I}(f)) = (f_*(\Omega_Y))(X) \cong \mathcal{O}(Y)$, so we have only to verify that the frame homomorphism $s: \mathcal{O}(X) \to (f_*(\Omega_Y))(X)$ coincides with f^*. But this is easy, since if we identify $f_*(\Omega_Y)(U)$ with $\{V \in \mathcal{O}(Y) \mid V \leq f^*(U)\}$ then its top element $\top_U$ is $f^*(U)$, and σ_X^U is identified with the inclusion map. Once again, the naturality of this isomorphism $\mathcal{E}\mathcal{I}(f) \cong f$ is easy to verify. □

We recall from 1.3.11 that $\mathbf{Sh}(X)$ itself is equivalent to a full subcategory of $\mathbf{Loc}/X$, namely $\mathbf{LH}/X$; so Theorem 1.6.3 identifies it with a full subcategory of $\mathbf{Loc}(\mathbf{Sh}(X))$. Of course, the objects of this subcategory are the *discrete* internal locales (that is, those whose frames are power objects Ω^A): if $f: E \to X$ is a local homeomorphism, then $\mathbf{Sh}(E) \simeq \mathbf{Sh}(X)/\Gamma(f)$, where $\Gamma(f)$ is the sheaf of sections of f, and so $f_*(\Omega_E) \cong f_*f^*(\Omega_X) \cong \Omega_X^{\Gamma(f)}$ by A2.3.2 and A1.5.2. More generally, it is straightforward to check that the functors $\mathcal{I}$ and $\mathcal{E}$ both preserve local homeomorphisms, and so we obtain

Scholium 1.6.4 *For any locale morphism $f: Y \to X$, the topos* $\mathbf{Sh}(Y)$ *is equivalent (as a topos defined over* $\mathbf{Sh}(X)$*) to the topos of sheaves on the internal locale $\mathcal{I}(f)$ in* $\mathbf{Sh}(X)$. □

We say an internal locale Z is *spatial* if its frame $\mathcal{O}(Z)$ is (isomorphic to) a subframe of a power object Ω^A – equivalently, if there exists a locale epimorphism from a discrete locale A to Z. Translating this via the equivalence of 1.6.3, we obtain

Proposition 1.6.5 *Let $f: Y \to X$ be a locale morphism. The following conditions are equivalent:*

(i) *f corresponds to a spatial internal locale in* $\mathbf{Sh}(X)$.

(ii) *There exists a locale epimorphism $h: E \to Y$ such that fh is a local homeomorphism.*

(iii) *The canonical map $\Delta\Theta\Gamma(f) \to f$ (the counit of the adjunction of* 1.3.12(ii)*) is an epimorphism.*

(iv) *The family of all partial sections of f over open sublocales of X is epimorphic.*

Proof (i)$\Leftrightarrow$(ii): Since $\mathcal{I}$ and $\mathcal{E}$ form an equivalence of categories, they preserve and reflect epimorphisms (and so does the forgetful functor $\mathbf{Loc}/X \to \mathbf{Loc}$). So this is immediate from the identification of discrete internal locales in $\mathbf{Sh}(X)$ with local homeomorphisms over X.

(ii)$\Leftrightarrow$(iii) since any epimorphism in $\mathbf{Loc}/X$ from a local homeomorphism to f must factor through the counit.

(iii) $\Leftrightarrow$ (iv): By definition, $\Delta\Theta\Gamma(f)$ is the colimit in $\mathbf{Loc}/X$ of a diagram whose vertices are open sublocales of X, and under which the partial sections of f form a cone. So the family of partial sections is epimorphic iff the induced morphism from the colimit (i.e. the counit map) is epic. □

Corollary 1.6.6 *Suppose $f: Y \to X$ corresponds to a spatial internal locale in* $\mathbf{Sh}(X)$. *Then*

(i) *f is an open map.*

(ii) *If X is spatial, so is Y.*

Proof (i) It is clear from the definition of open maps that if fh is open and h is epimorphic, then f is open. So this follows from 1.6.5(ii) and the fact that local homeomorphisms are open maps.

(ii) similarly follows from 1.3.2(v) and the fact that an epimorphic image of a spatial locale is spatial. □

From 1.6.6(ii), we see that if X is a sober space then the category $\mathbf{Sob}(\mathbf{Sh}(X))$ of internal sober spaces in $\mathbf{Sh}(X)$ (equivalent to internal spatial locales) may be identified with a full subcategory of $\mathbf{Sob}/X$. But 1.6.6(i) shows that it is by no means the whole of this category. In fact, for T_D-spaces, the spaces over X which correspond to spatial internal locales are characterized by a familiar condition:

Corollary 1.6.7 *Let $f: Y \to X$ be a continuous map of sober spaces. If Y is a T_D-space, then f corresponds to a spatial internal locale in* $\mathbf{Sh}(X)$ *iff it is a* submersion, *i.e. every point of Y lies in the image of a continuous section of f over some open subset of X.*

Proof This is immediate from 1.6.5(iv), since a family of maps to a sober T_D-space is epimorphic in $\mathbf{Loc}$ iff it is surjective (cf. the remarks after 1.5.2). □

Example 1.6.8 To fulfil a promise made after Lemma 1.2.4, we now give an example to show that that result can fail in a non-Boolean topos. Let X be a Hausdorff space (in $\mathbf{Set}$) which is not discrete, and let $\tilde{X}$ be the internal locale in $\mathbf{Sh}(X)$ corresponding to the projection $\pi_1: X \times X \to X$. π_1 is a submersion, since any point (x, y) of $X \times X$ is in the image of the (global) section $t \mapsto (t, y)$; so $\tilde{X}$ is a spatial internal locale in $\mathbf{Sh}(X)$. Indeed, since the identity map $X \times X_{dp} \to X \times X$ (where X_{dp} denotes X retopologized with the discrete

topology, cf. the remarks after 1.2.13) is a continuous surjection, we can regard $\mathcal{O}(\tilde{X})$ as a topology on the sheaf $\check{X} = \Gamma(\pi_1 \colon X \times X_{dp} \to X)$, i.e. the sheaf whose sections over U are all locally constant functions $U \to X$. And since X is (externally) Hausdorff, it is not hard to verify that this topology makes $\check{X}$ into an internal Hausdorff space in $\mathbf{Sh}(X)$. But it is not sober: its sobrification is (the 'sheaf of points' of) $\tilde{X}$, i.e. the sheaf $\Gamma(\pi_1 \colon X \times X \to X)$ whose sections over U are all continuous functions $U \to X$. And this is strictly larger than $\check{X}$, since the identity function $X \to X$ is not locally constant.

We may also exploit Theorem 1.6.3 along the lines foreshadowed at the beginning of Section C1.4, by establishing the equivalence of various topological properties of internal locales in $\mathbf{Sh}(X)$ and properties of the locale maps $Y \to X$ which correspond to them. However, we shall postpone this line of development until Chapter C3, when we shall investigate it in some particular cases.

For future reference, we shall need a description of internal frames in a topos of the form $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, where $\mathcal{C}$ is a small cartesian category, similar to the description of internal frames in $\mathbf{Sh}(X)$ which we gave earlier in this section. For this, it is convenient to proceed by way of complete join-semilattices.

Lemma 1.6.9 *Let $\mathcal{C}$ be a small cartesian category, and let $L \colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Poset}$ be a functor, considered as an internal ordered object in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.*

(i) *L is an internal complete join-semilattice in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ iff it takes values in the category **CLat** of complete lattices (i.e. $L(U)$ is a complete (semi-) lattice for each object U of $\mathcal{C}$, and $L(a)$ preserves arbitrary joins and meets – equivalently, has adjoints on both sides – for each morphism a), and the Beck–Chevalley condition holds, i.e., for each pullback square*

$$\begin{tikzcd} U \arrow[r, "a"] \arrow[d, "b"] & V \arrow[d, "c"] \\ W \arrow[r, "d"] & X \end{tikzcd}$$

in $\mathcal{C}$, the square

$$\begin{tikzcd} L(V) \arrow[r, "L(a)"] \arrow[d, "\Sigma_c"] & L(U) \arrow[d, "\Sigma_b"] \\ L(X) \arrow[r, "L(d)"] & L(W) \end{tikzcd}$$

(where Σ_b and Σ_c are the left adjoints of $L(b)$ and $L(c)$) commutes.

(ii) *L is an internal frame in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ iff it satisfies the conditions in* (i), *and additionally each $L(U)$ is a frame and the Frobenius reciprocity condition*

$$\Sigma_a(L(a)(u) \wedge v) = u \wedge \Sigma_a(v)$$

holds for all $a\colon V \to U$ in $\mathcal{C}$, $u \in L(U)$ and $v \in L(V)$.

Proof (i) First, suppose that L is an internal complete semilattice; equivalently, that the $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$-indexed category $[L]$ defined as in B2.3.3 is cocomplete. By B2.4.8 (or by the remark after B2.3.9), it is also complete. To see that the individual posets $L(U)$ are necessarily complete and cocomplete, we observe that $F \mapsto F(U)$ is the direct image of a pre-geometric morphism; specifically, it is the composite

$$[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{Y(U)^*} [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]/Y(U) \xrightarrow{\gamma_*} \mathbf{Set}$$

where Y denotes the Yoneda embedding and γ_* is the direct image of the unique geometric morphism. So by B2.3.7 it preserves (co)completeness of posets. The Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ preserves pullbacks, so by B1.4.9 this implies that the $\mathcal{C}$-indexed category obtained by restricting $[L]$ to representables is also complete and cocomplete. But this is just L itself regarded as a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Cat}$; so we deduce that the transition maps $L(a)$ have left and right adjoints, and that the Beck–Chevalley conditions hold. (Of course, either of them implies the other, as we observed in B1.4.6.)

For the converse, we use the characterization of cocomplete internal posets provided by B2.3.9(ii): that is, we construct a join map $\bigvee\colon PL \to L$ left adjoint to the principal-ideal map $\downarrow\colon L \to PL$. We recall that $PL(U)$ may be identified with the set of all subfunctors of $L \times \mathcal{C}(-, U)$; and $\downarrow_U$ sends an element u to the subfunctor whose value at V is the set of pairs (v, a) such that $v \leq L(a)(u)$. Given an arbitrary subfunctor $F \rightarrowtail L \times \mathcal{C}(-, U)$, we define $\bigvee_U(F)$ to be the join, over all $V \in \mathrm{ob}\,\mathcal{C}$ and all $(v, a) \in F(V)$, of the elements $\Sigma_a(v)$ of $L(U)$. To see that this defines a natural transformation $PL \to L$, we use the Beck–Chevalley condition: given $b\colon W \to U$ in $\mathcal{C}$ and $(v, a) \in F(V)$, we have $L(b)\Sigma_a(v) = \Sigma_c L(d)(v)$, where

$$\begin{tikzcd} X \arrow[r, "c"] \arrow[d, "d"] & W \arrow[d, "b"] \\ V \arrow[r, "a"] & U \end{tikzcd}$$

is a pullback in $\mathcal{C}$, and the pair $(L(d)(v), c)$ belongs to $PL(b)(F)(X)$. So naturality follows from the fact that $L(b)$ preserves joins. It is also clear that $\bigvee_U \downarrow_U (u) \leq u$ for all $u \in L(U)$, since the inequality $v \leq L(a)(u)$ is equivalent to $\Sigma_a(v) \leq u$ (in fact we actually have $\bigvee_U \downarrow_U (u) = u$, since the pair $(u, 1_U)$

belongs to $\downarrow_U (u)(U)$); and $\downarrow_U \bigvee_U(F) \supseteq F$ for all $F \in PL(U)$, by a similar argument. So we have the required left adjoint.

(ii) Suppose L is an internal frame. Then by 1.1.2 it is an internal Heyting algebra, so each $L(U)$ is a Heyting algebra and hence a frame. Also, each $L(a)$ is a Heyting algebra homomorphism, so by A1.5.8 the Frobenius reciprocity law holds. Conversely, if each $L(U)$ is a frame and Frobenius reciprocity holds, then L is an internal Heyting algebra and hence a frame. $\square$

We may easily extend 1.6.9(ii) from functor categories $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ to toposes $\mathbf{Sh}(\mathcal{C}, T)$ of sheaves for a coverage on $\mathcal{C}$. We recall from B2.3.10(ii) that, if j is a local operator on a topos $\mathcal{E}$, then the complete internal posets in $\mathbf{sh}_j(\mathcal{E})$ are just those complete internal posets in $\mathcal{E}$ whose underlying objects are j-sheaves; since the same is clearly true for Heyting algebras (by D1.2.13, if you insist), we deduce that the internal frames in $\mathbf{sh}_j(\mathcal{E})$ are simply those internal frames in $\mathcal{E}$ whose underlying objects are sheaves. Also, by A4.3.5 and A4.4.8, we know that every coverage on a small category $\mathcal{C}$ corresponds to a local operator on $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. So we may immediately deduce

Corollary 1.6.10 *Let $(\mathcal{C}, T)$ be a small site whose underlying category $\mathcal{C}$ is cartesian. Then internal frames in $\mathbf{Sh}(\mathcal{C}, T)$ may be identified with functors $L\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{CLat}$ such that each $L(U)$ is a frame, the underlying set-valued functor of L is a T-sheaf, and L satisfies the Beck–Chevalley condition of* 1.6.9(i) *and the Frobenius reciprocity condition of* 1.6.9(ii). $\square$

Suggestions for further reading: Fourman & Grayson [363], Joyal & Tierney [560], Moerdijk [825], Stout [1122].

C2

SHEAVES ON A SITE

C2.1 Sites and coverages

We have seen in the last chapter how we may regard many familiar topological properties (of spaces or locales) as properties of sheaf toposes. This in itself would not be a significant gain, were it not for the fact that there are other contexts in mathematics where 'topological intuition' seems to be present, and which cannot be adequately described by either spaces or locales – but these contexts *do* give rise to toposes.

Such contexts first arose in algebraic geometry, in the study of cohomology theories such as étale cohomology (see Chapter E3): here one has a cohomology theory living on something which resembles a space in that it possesses 'localizations' and 'coverings', but which has an additional feature, not present in topology, that its 'points' have nontrivial automorphism groups which form part of the structure of the 'space'. It was A. Grothendieck [36] who first saw how to give a unified description of such contexts, via the notion of a *site*.

Grothendieck's idea was that, in the definition of a sheaf on a space (or locale) X, the fact that the category $\mathcal{O}(X)$ on which the presheaves are defined is a frame is incidental: all one needs is a good notion of 'covering' in the category, in order to be able to formulate the sheaf axiom. Grothendieck used the name 'topology' for this notion of covering; but this name, and the rather rigid axiomatization of the notion which he formulated, has unhelpfully tended to obscure the real magnitude of the generalization involved in passing from frames to sites (categories equipped with a notion of covering). In particular, the fact that in a frame the coverings are determined by the structure of the category (they are exactly the colimit cones) has no parallel in the more general context, where the notion of covering is something arbitrarily imposed on a given category. Thus we have chosen to adopt the name 'coverage', which carries less emotional baggage around with it than 'topology', for the structure which is needed in order to define sheaves.

A further fruitful source of possible misunderstanding in Grothendieck's account [36] is that his notion of 'topology' is defined by three closure conditions of which one plays a very different rôle from the other two: one (in fact the second, in the order in which they are usually stated) is essential to the definition of the category of sheaves (cf. 2.1.13 below), whereas the other two

are 'saturation conditions' which might as well be there because one can always close off under them without affecting the notion of sheaf. In our account, we shall (we hope) keep the distinction between the former and latter conditions clear; in fact we start off by assuming only the former.

The definition of a coverage has already been given in A2.1.9, but we shall repeat it here – with one minor change: in Section A2.1 we considered coverages only on small categories, but here we shall not assume anything (for the time being, at least) about the size of the underlying category. (We also make one notational change: we shall denote objects of the underlying category by letters like $U, V, W, \ldots$ rather than $A, B, C, \ldots$, partly to maintain continuity with the previous chapter and partly because we want to reserve $A, B, C, \ldots$ for objects of the topos of sheaves.)

Definition 2.1.1 Let $\mathcal{C}$ be a category. By a *coverage* on $\mathcal{C}$, we mean a function assigning to each object U of $\mathcal{C}$ a collection $T(U)$ of families $(f_i : U_i \to U \mid i \in I)$ of morphisms with common codomain U (called *T-covering families*), such that

(C) If $(f_i : U_i \to U \mid i \in I)$ is a T-covering family and $g : V \to U$ is any morphism with codomain U, there exists a T-covering family $(h_j : V_j \to V \mid j \in J)$ such that each gh_j factors through some f_i.

A *site* is a category equipped with a coverage. By a *small site*, we mean one whose underlying category is small.

As we have already indicated, the purpose of introducing sites is in order to be able to define sheaves. Once again, we restate the definition from Section A2.1; for convenience, we also recall the definition of separated functor from A2.6.4(d).

Definition 2.1.2 Let $\mathcal{C}$ be a category. We say a functor $A : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ *satisfies the sheaf axiom* for a family of morphisms $(f_i : U_i \to U \mid i \in I)$ if, whenever we are given a family of elements $s_i \in A(U_i)$ which are *compatible* in the sense that, whenever $g : V \to U_i$ and $h : V \to U_j$ satisfy $f_i g = f_j h$ (here i and j need not be distinct), we have $A(g)(s_i) = A(h)(s_j)$, then there exists a unique $s \in A(U)$ such that $A(f_i)(s) = s_i$ for each $i \in I$. We say A is *separated* for $(f_i \mid i \in I)$ if it satisfies the above condition with 'a unique' replaced by 'at most one'. If T is a coverage on $\mathcal{C}$, we say A is a *T-sheaf* (resp. *T-separated*) if it satisfies the sheaf axiom (resp. is separated) for every T-covering family. And we write $\mathbf{Sh}(\mathcal{C}, T)$ (resp. $\mathbf{Sep}(\mathcal{C}, T)$) for the full subcategory of the functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ whose objects are T-sheaves (resp. T-separated functors).

The collection of all families of morphisms for which a given functor A satisfies the sheaf axiom has a number of important closure properties; so there is no harm in adding these closure properties to the condition (C) which we have assumed in 2.1.1. We begin with the simplest one: recall that a *sieve* on an object U of $\mathcal{C}$ is a family R of morphisms with codomain U which is a right ideal in $\mathcal{C}$, i.e. $f \in R$ implies $fg \in R$ whenever the composite fg is defined. In the terminology introduced in A4.5.2 and A4.5.5, sieves on an object U of $\mathcal{C}$ are

sieves in the slice category $\mathcal{C}/U$. (If $\mathcal{C}$ is locally small – which it generally will be from now on – then sieves on U correspond bijectively to subobjects in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ of the representable functor $\mathcal{C}(-, U)$, the correspondence sending a sieve R to the functor $V \mapsto \{f \in R \mid \mathrm{dom}\, f = V\}$.) Any family $(f_i : U_i \to U \mid i \in I)$ generates a sieve on U, consisting of all the morphisms with codomain U which factor through at least one f_i.

Lemma 2.1.3 *Let $(f_i : U_i \to U \mid i \in I)$ be a family of morphisms of $\mathcal{C}$ with common codomain, and $A : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ a functor. Then A satisfies the sheaf axiom (resp. is separated) for $(f_i \mid i \in I)$ iff it satisfies the sheaf axiom (resp. is separated) for the sieve generated by this family.*

Proof From the very definition of a compatible family, it is easy to see that any compatible family with respect to $(f_i \mid i \in I)$ extends uniquely to a compatible family with respect to the sieve generated by the f_i. So this is immediate. □

It is clear that a collection T of families of morphisms satisfies condition (C) iff the same is true for the collection $\overline{T}$ consisting of all the sieves generated by members of T; so we may replace any site $(\mathcal{C}, T)$ by a site $(\mathcal{C}, \overline{T})$ which has the same sheaves, in which all the covering families are sieves. We shall call a coverage *sifted* if it has this latter property. Note that the definition of compatibility with respect to a family of morphisms can be simplified when the family is a sieve: $(s_f \mid f \in R)$ is compatible with respect to a sieve R iff, whenever $f \in R$ and $\mathrm{dom}\, f = \mathrm{cod}\, g$, we have $s_{fg} = A(g)(s_f)$. A further simplification is possible in the case when $\mathcal{C}$ is locally small, so that we can identify sieves with subfunctors of representable functors: in this case, a compatible family $(s_f \mid f \in R)$ of elements of A is exactly (the image of) a natural transformation $R \to A$. Bearing in mind the Yoneda Lemma, which says that elements of $A(U)$ correspond to natural transformations $\mathcal{C}(-, U) \to A$, we thus obtain

Lemma 2.1.4 *Let $\mathcal{C}$ be a locally small category, and R a sieve on an object U of $\mathcal{C}$. Then a functor $A : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ satisfies the sheaf axiom (resp. is separated) for R iff each natural transformation $R \to A$ factors uniquely (resp. in at most one way) through the inclusion $R \rightarrowtail \mathcal{C}(-, U)$.* □

One further advantage of restricting our attention to sifted coverages is that the statement of axiom (C) can be simplified. Given a sieve R on U and a morphism $f : V \to U$ we define $f^*(R)$ to be the family (which is clearly a sieve on V) of all morphisms g with codomain V such that $fg \in R$. (In the case when $\mathcal{C}$ is locally small and we identify sieves with subfunctors of representable

functors, this is simply the pullback

$$
\begin{tikzcd}
f^*(R) \arrow[r] \arrow[d, tail] & R \arrow[d, tail] \\
\mathcal{C}(-,V) \arrow[r] & \mathcal{C}(-,U)
\end{tikzcd}
$$

where the bottom edge of the square is the natural transformation induced by f.) Then condition (C), for a sifted coverage T, is equivalent to saying that if $R \in T$ then $f^*(R)$ contains a sieve in T for every morphism f whose codomain matches that of R.

However, there are also disadvantages in passing from arbitrary covering families to sieves. The principal one is that, in many particular cases (see the list of examples of sites in A2.1.11), the covering families may be easier to specify explicitly than the sieves which they generate (for example, the generating families may be finite); and the verification that particular functors are sheaves also tends to be easier when we work with the generating families, at least in the case when the underlying category $\mathcal{C}$ has pullbacks. For this reason, we shall state the next three lemmas in two different versions, one for generating families and the other for sieves. (They also have alternative versions for separated functors rather than sheaves, but we shall leave the reader to supply these.)

Lemma 2.1.5 *Let U be an object of $\mathcal{C}$, and $A \colon \mathcal{C}^{\mathrm{op}} \to$ **Set** a functor.*

(i) *A satisfies the sheaf axiom for the singleton family (1_U).*

(ii) *A satisfies the sheaf axiom for the sieve M_U of all morphisms with codomain U.*

Proof The two statements are equivalent, since the family (1_U) generates the sieve M_U. And (i) is immediate from Definition 2.1.2. □

Lemma 2.1.6 *Let $A \colon \mathcal{C}^{\mathrm{op}} \to$ **Set** be a sheaf for a coverage T on $\mathcal{C}$.*

(i) *Let $(f_i \colon U_i \to U \mid i \in I)$ be a T-covering family, and $(g_j \colon V_j \to U \mid j \in J)$ a family of morphisms with the same codomain, such that every f_i factors through at least one g_j. Then A satisfies the sheaf axiom for $(g_j \mid j \in J)$.*

(ii) *If T is a sifted coverage, then A satisfies the sheaf axiom for any sieve which contains a T-covering sieve.*

Proof It suffices to prove (i), since (ii) is a direct translation of it into 'sifted' language. For (i), let $(t_j \in A(V_j) \mid j \in J)$ be a compatible family with respect to the g_j. For each $i \in I$, we can choose a factorization $f_i = g_{j(i)}h_i$ of f_i through some g_j, and define $s_i = A(h_i)(t_{j(i)}) \in A(U_i)$. The compatibility of the t_j ensures that s_i is independent of the choice of $j(i)$ and h_i, and that the s_i form a compatible family with respect to the f_i; so there is a unique $s \in A(U)$

such that $A(f_i)(s) = s_i$ for each $i \in I$. This is clearly the unique element of $A(U)$ which could possibly satisfy $A(g_j)(s) = t_j$ for all $j \in J$; to see that it actually does so (for a particular j), we need to observe that if $k: W \to V_j$ is a morphism such that $g_j k$ factors through some f_i (say $g_j k = f_i l$), then

$$A(k)A(g_j)(s) = A(l)A(f_i)(s) = A(l)(s_i) = A(l)A(h_i)(t_{j(i)}) = A(k)(t_j)$$

(where the last step uses the compatibility of the t_j). But by condition (C) there is a T-covering family of morphisms k with this property, and so the result follows since A is a T-sheaf. □

Lemma 2.1.7 *Let A be a functor $\mathcal{C}^{\text{op}} \to$* **Set**.

(i) *Let $(f_i: U_i \to U \mid i \in I)$ be a family for which A satisfies the sheaf axiom, and suppose, for each $i \in I$, there is a family $(h_{ij}: U_{ij} \to U_i \mid j \in J_i)$ for which A also satisfies the sheaf axiom. Then A satisfies the sheaf axiom for the family of all composites $(f_i h_{ij} \mid i \in I, j \in J_i)$.*

(ii) *Suppose T is a sifted coverage for which A is a sheaf. Let R be a T-covering sieve on an object U of $\mathcal{C}$, and S a sieve on U such that $f^*(S)$ is a T-covering sieve on* dom f *for each $f \in R$. Then A satisfies the sheaf axiom for S.*

Proof (i) Suppose given a compatible family $(s_{ij} \in A(U_{ij}) \mid i \in I, j \in J_i)$. For each fixed $i \in I$, the family $(s_{ij} \mid j \in J_i)$ is compatible with respect to the family $(h_{ij} \mid J \in J_i)$, and so defines a unique element $s_i \in A(U_i)$ such that $A(h_{ij})(s_i) = s_{ij}$ for all $j \in J_i$. Further, the compatibility of the s_{ij} ensures that $(s_i \mid i \in I)$ is a compatible family relative to the f_i. Since A satisfies the sheaf axiom for the f_i, there is a unique $s \in A(U)$ satisfying $A(f_i)(s) = s_i$ for all i.

(ii) The second version of the statement is not quite a simple translation of the first into the language of sieves, since S may be strictly larger than the sieve of all composites fh with $f \in R$ and $h \in f^*(S)$ – there may be morphisms in S which do not factor through any member of R. But Lemma 2.1.6(ii) provides the extra information we need. □

The next definition, like the last three lemmas, could be given separately in 'unsifted' and 'sifted' versions, but from now on we shall leave to the reader the task of formulating the former.

Definition 2.1.8 We define a *Grothendieck coverage* on a category $\mathcal{C}$ to be a function assigning to each object U of $\mathcal{C}$ a collection $T(U)$ of sieves on U, satisfying the closure property (C) of 2.1.1 and also

(M) For any object U, the maximal sieve M_U is in $T(U)$.

(L) If $R \in T(U)$ and S is another sieve on U such that, for each $f \in R$, the sieve $f^*(S)$ belongs to $T(\text{dom } f)$, then S also belongs to $T(U)$.

The condition (L) is sometimes referred to as the *local character condition.* We note that, if a sifted coverage T satisfies both (L) and (M), then any sieve which contains a T-covering sieve must be T-covering; for if $R \subseteq S$, then $f^*(S) = M_{\mathrm{dom}\, f}$ for each $f \in R$. Hence, in the presence of (L) and (M), (C) may be replaced by the simpler condition

(C') If $R \in T(U)$ and g is any morphism with codomain U, then $g^*(R) \in T(\mathrm{dom}\, g)$.

(This is the form in which the definition of 'Grothendieck topology' is usually presented.) One further closure property which we could have demanded is that $T(U)$ should be closed under finite intersections (and so form a filter in the lattice of subfunctors of $\mathcal{C}(-, U)$); the fact that the sieves for which a given A satisfies the sheaf axiom are closed under finite intersections can be proved by similar means to the last three lemmas. However, there is no need to do so explicitly, since this closure property follows from (L) and (C'): for any $f \in R$, we have $f^*(R \cap S) = f^*(S)$.

From Lemmas 2.1.5, 2.1.6 and 2.1.7, we immediately deduce

Proposition 2.1.9 *For any coverage T on a category $\mathcal{C}$, there is a Grothendieck coverage $\widetilde{T}$ having the same sheaves as T.*

Proof First replace T by a sifted coverage $\overline{T}$ as in 2.1.3, and then take $\widetilde{T}$ to be the intersection of all Grothendieck coverages which contain $\overline{T}$. It is clear from the nature of the closure conditions (L) and (M) that $\widetilde{T}$ is itself a Grothendieck coverage. Moreover, if A is any T-sheaf then it is also a $\overline{T}$-sheaf by 2.1.3; and so it follows from the last three lemmas that the collection of all sieves R such that A satisfies the sheaf axiom for every $f^*(R)$ (as f ranges over all morphisms with the same codomain as the members of R) is a Grothendieck coverage containing $\overline{T}$. So A is a sheaf for $\widetilde{T}$. The converse is immediate. □

In the case when $\mathcal{C}$ is small, $\widetilde{T}$ can also be constructed from $\overline{T}$ by a transfinite iteration: first adjoin all the maximal sieves M_A to $\overline{T}$, and then repeatedly adjoin those sieves whose presence is required by (L), until the process converges. However, if $\mathcal{C}$ is large, there may be a proper class of distinct sieves on a given object of $\mathcal{C}$, and so there is no guarantee that this process will converge.

For the rest of this and the next two sections, we shall assume that all coverages are Grothendieck coverages, unless explicitly stated otherwise. We shall also tend to denote Grothendieck coverages by letters like J and K rather than T: in addition to drawing attention to the fact that a particular coverage is a Grothendieck coverage, this notation is consistent with that which we used in Sections A4.4 and A4.5, in the following sense.

Proposition 2.1.10 *Let $\mathcal{C}$ be a small category, and J a function assigning to each object U of $\mathcal{C}$ a set $J(U)$ of sieves on U. Then J satisfies condition (C') iff it is a subfunctor of the subobject classifier Ω of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. Moreover, if this*

condition is satisfied, then J is a Grothendieck coverage iff the classifying map $j\colon \Omega \to \Omega$ of $J \rightarrowtail \Omega$ is a local operator in the sense of A4.4.1.

Proof The first assertion is immediate from the definition of Ω in A1.6.6. Given this, the natural transformation $j\colon \Omega \to \Omega$ corresponding to J is easily seen to be defined by

$$j_U(R) = \{f\colon V \to U \mid f^*R \in J(V)\},$$

so it is clear that J satisfies (M) iff $1_U \in j_U(M_U)$ for all U, iff $j\top = \top$. If condition (L) also holds, then we have $j_U(R) \in J(U)$ iff $R \in J(U)$, so that $jj = j$; and from the fact, noted above, that the set of J-covering sieves is closed under finite intersections we similarly deduce that j satisfies the third condition of A4.4.1. Conversely, if j satisfies the second and third conditions of A4.4.1, and R and S are two sieves on U satisfying the hypotheses of (L), then we have $R \subseteq j_U(S)$, whence $j_U(S) = j_U j_U(S) \supseteq j_U(R) = M_U$, so $S \in J(U)$. □

We may now deduce (at least for small categories) the result that a Grothendieck coverage is determined by its sheaves; this is sometimes known as the 'little Giraud theorem'.

Corollary 2.1.11 *For a small category $\mathcal{C}$, the assignment $J \mapsto$ **Sh**$(\mathcal{C}, J)$ is a bijection from the set of Grothendieck coverages on $\mathcal{C}$ to the class of reflective subcategories of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ with cartesian reflector.*

Proof In A4.4.8 we established a bijection between such subcategories and local operators on $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. So the only remaining ingredient is to show that if J and j are related as in 2.1.10 then $\mathbf{sh}_j([\mathcal{C}^{\mathrm{op}}, \mathbf{Set}])$ coincides with $\mathbf{Sh}(\mathcal{C}, J)$. But this was essentially done in A4.3.5: for the operator d on subobjects defined there is easily seen to coincide with that obtained by composing classifying maps with j (and in particular it is always idempotent if J is a Grothendieck coverage). □

We shall extend 2.1.11 to non-small categories in 2.2.5 below.

We saw in the proof of 2.1.9 that, for a given functor $A\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, there is a unique largest sifted coverage for which A is a sheaf, and it is a Grothendieck coverage. By intersecting such coverages, we may construct the largest coverage for which each of a given family of functors is a sheaf. In particular, when $\mathcal{C}$ is locally small, we define the *canonical coverage* on $\mathcal{C}$ to be the largest sifted coverage for which all representable functors are sheaves.

We say a sieve R on U is *effective-epimorphic* if it forms a colimit cone under the (large!) diagram consisting of the domains of all the morphisms in R, and all the morphisms over U between them. And we say R is *universally effective-epimorphic* if $f^*(R)$ is effective-epimorphic for every f with codomain U. It is easy to see that a sieve is effective-epimorphic iff every representable functor satisfies the sheaf axiom for it; hence the universally effective-epimorphic sieves

are exactly the covering sieves for the canonical coverage. We say a coverage is *subcanonical* if all its covering sieves are effective-epimorphic, i.e. if it is contained in the canonical coverage; we say a site $(\mathcal{C}, J)$ is *standard* if $\mathcal{C}$ is small and cartesian, and J is subcanonical. We shall also call a coverage *separated* if all its covering sieves are epimorphic families; clearly, this is equivalent to saying that all the representable functors are separated in the sense of 2.1.2.

We shall see in the next section that there is no loss of generality (at least if we are interested not in the sites themselves, but only in their categories of sheaves) in restricting our attention to standard sites.

Examples 2.1.12 (*a*) The empty sieve on an object U of $\mathcal{C}$ is effective-epimorphic iff U is an initial object; it is universally effective-epimorphic iff U is a strict initial object (cf. A1.4.1). (And it is merely epimorphic iff U is a quasi-initial object as defined before A1.5.14.)

(*b*) Let R be the sieve generated by a single morphism $f\colon V \to U$. Then R is effective-epimorphic iff any morphism $g\colon V \to W$ which satisfies $gh = gk$ whenever $fh = fk$ factors uniquely through f. If f has a kernel-pair $(a, b)\colon K \rightrightarrows V$, then g satisfies the above condition iff $ga = gb$; so R is effective-epimorphic iff f is a coequalizer of its kernel-pair, i.e. iff f is regular epic. If $\mathcal{C}$ has pullbacks, then R is universally effective-epimorphic iff every pullback of f is regular epic; in particular, we deduce that *in a regular category, the sieve generated by a single morphism f is universally effective-epimorphic iff f is a cover* (a fact which we used in Example A2.1.11(*a*)). If $\mathcal{C}$ is a regular category, we define the *regular coverage* on $\mathcal{C}$ to consist of all sieves which contain a cover (in the sense of A1.3.2); it is easy to see that this is in fact a Grothendieck coverage, and if $\mathcal{C}$ is small then the resulting site is standard.

(*c*) We saw in A2.1.11(*h*) that the collection of all singleton families of morphisms defines a coverage on $\mathcal{C}$ iff $\mathcal{C}$ satisfies the 'right Ore condition' that any pair of morphisms with common codomain may be embedded in a commutative square. The Grothendieck coverage generated by this coverage as in 2.1.9 clearly consists of all inhabited (= nonempty) sieves on objects of $\mathcal{C}$. Can it be the canonical coverage? Clearly, this will be the case if every sieve generated by a singleton is effective-epimorphic, and the empty sieve is not universally effective-epimorphic on any object. So, combining (*a*) and (*b*) above, we deduce that *if $\mathcal{C}$ is a category satisfying the right Ore condition, in which every morphism is regular epic and no object is strict initial, then the canonical coverage on $\mathcal{C}$ consists of all inhabited sieves.* Examples of such categories are the category of (finite) sets and epimorphisms, and the dual of the category of (finite) sets and monomorphisms; we shall meet others in 3.5.9 below.

(*d*) Let $\mathcal{C}$ be a coherent category, and let $(f_i\colon V_i \to U \mid 1 \leq i \leq n)$ be a finite covering family, i.e. one such that the union of the images $U_i \rightarrowtail U$ of the f_i is the whole of U. Then the sieve R generated by the f_i is effective-epimorphic (and hence universally effective-epimorphic, since finite covering families are stable under pullback in $\mathcal{C}$); for if $(g_i\colon V_i \to W \mid 1 \leq i \leq n)$ is a compatible family of

elements of $\mathcal{C}(-, W)$ relative to the f_i, then each g_i factors through $V_i \twoheadrightarrow U_i$ (since it coequalizes the kernel-pair of f_i) and the induced morphisms $U_i \to W$ can be combined, by repeated use of A1.4.3, into a unique morphism $U \to W$. (Once again, this argument was sketched before in A2.1.11(b).) We define the *coherent coverage* on a coherent category $\mathcal{C}$ to consist of all sieves which contain a finite covering family; by the remarks above, this coverage is subcanonical. If $\mathcal{C}$ is a pretopos, then finite covering families are the same thing as finite (jointly) epimorphic families, by A1.4.9; the coherent coverage on a pretopos is sometimes called the *pre-canonical coverage*.

(e) Similarly, if $\mathcal{C}$ is a geometric category (resp. an ∞-pretopos), we may show that sieves generated by small covering families (resp. small epimorphic families) are universally effective-epimorphic. There is a converse result if $\mathcal{C}$ has a separating set of objects, as defined in Section A1.2: for then any epimorphic sieve (in particular, any effective-epimorphic sieve) on an object U must contain an epimorphic family of morphisms $V \to U$ whose domains are all in the separating set, and there is only a set of such morphisms since a geometric category is locally small. Thus *in an ∞-pretopos with a separating set of objects, the canonical coverage consists precisely of those sieves which contain small epimorphic families.* This result will be of considerable importance in the next section.

Before leaving this section, we give an example to emphasize the difference already mentioned between condition (C) and the other closure conditions that we have introduced in the course of the section: if we omit the former from our definition of a coverage, we arrive at a notion of 'category of sheaves' very different from that which we wish to study. For simplicity, we formulate the definitions in this example in terms of posets, but they can in fact be considered for more general categories; cf. [540].

Example 2.1.13 Given a poset P, we define an *interval* in P to be a (nonempty) subset of the form

$$[a, b] = \{x \in P \mid a \leq x \leq b\}$$

with $a \leq b$ in P. We write $\mathrm{Int}(P)$ for the set of all intervals in P, ordered by inclusion. Int is a functor **Poset** $\to$ **Poset**: given an order-preserving map $f: P \to Q$, we define $\mathrm{Int}(f)([a, b]) = [f(a), f(b)]$. We say that $f: P \to Q$ is an *interval fibration* if, for every $a \leq b$ in P, the restriction of f to $[a, b]$ is an order-isomorphism $[a, b] \to [f(a), f(b)]$; equivalently if, given $u \in Q$ with $f(a) \leq u \leq f(b)$, there exists a unique $c \in P$ with $a \leq c \leq b$ and $f(c) = u$. It is not hard to verify that f is an interval fibration if and only if $\mathrm{Int}(f): \mathrm{Int}(P) \to \mathrm{Int}(Q)$ is a discrete fibration in the sense of B1.3.11; that is, iff it corresponds to a functor $\mathrm{Int}(Q)^{\mathrm{op}} \to \mathbf{Set}$ via the Grothendieck construction of A1.1.7. Moreover, if $f: P \to Q$ and $f': P' \to Q$ are both interval fibrations, then order-preserving maps $P \to P'$ over Q (are necessarily interval fibrations, and) correspond bijectively to order-preserving maps $\mathrm{Int}(P) \to \mathrm{Int}(P')$ over $\mathrm{Int}(Q)$. Thus the category $\mathbf{IFib}/Q$ of interval fibrations over Q is equivalent to a full subcategory of $[\mathrm{Int}(Q)^{\mathrm{op}}, \mathbf{Set}]$.

An arbitrary discrete fibration $d: E \to \mathrm{Int}(Q)$ need not be of the form $\mathrm{Int}(f)$. However, if it is, then we may clearly reconstruct the domain P of f by taking its underlying set to be

$$\{a \in E \mid d(a) \text{ is a singleton }\},$$

with $a \leq b$ in P iff there exists $x \in E$ with $d(x) = [f(a), f(b)]$ and such that $a \leq x, b \leq x$ both hold in E. In general, this definition yields a relation on P which is reflexive and antisymmetric, but not necessarily transitive; for transitivity, what we need is precisely that the functor $\mathrm{Int}(Q)^{\mathrm{op}} \to \mathbf{Set}$ corresponding to d should satisfy the sheaf axiom for the family

$$([u,v] \subseteq [u,w] \ , \ [v,w] \subseteq [u,w])$$

whenever $u \leq v \leq w$ in Q. Thus we may conclude that $\mathbf{IFib}/Q$ is equivalent to '$\mathbf{Sh}(\mathrm{Int}(Q), T)$', where T is the set of all such families.

Now let Q be the four-element poset $\{0, u, v, 1\}$ with $0 \leq u \leq 1, 0 \leq v \leq 1$ but u and v incomparable. It is clear that, for this Q, the collection of covers T fails to satisfy condition (C): there is no T-covering family on $[v, v]$ which factors through the covering $([0, u] \subseteq [0, 1], [u, 1] \subseteq [0, 1])$ of $[0, 1]$. Moreover, in this case $\mathbf{IFib}/Q$ is very far from being a topos. To see this, note that monomorphisms in $\mathbf{IFib}/Q$ are just injective order-preserving maps, and so the subterminal objects of $\mathbf{IFib}/Q$ may be identified with those sub-posets of Q for which the inclusion is an interval fibration. Among these, we have the subsets $\{u\}$, $\{0, u\}$, $\{0, u, v\}$ and $\{u, 1\}$ as well as Q itself; but $\{0, u, 1\}$ is not a subterminal object, and so the join of $\{0, u\}$ and $\{u, 1\}$ in $\mathrm{Sub}_{\mathbf{IFib}/Q}(1)$ is the whole of Q. Thus the five subterminal objects mentioned above form a non-modular (and in particular non-distributive) lattice. It now follows from A1.4.2 that $\mathbf{IFib}/Q$ is not a coherent category, from A1.5.10 that it is not cartesian closed, and from A1.6.3 that it does not have a subobject classifier. (In fact it is not hard to show that $\mathbf{IFib}/Q$ is not even regular.)

Suggestions for further reading: Artin et al. [36], Bunge & Niefield [204], Johnstone [540], Pedicchio & Vermeulen [946].

C2.2 The topos of sheaves

In A2.1.10, we gave a sketch proof of the fact that the category of sheaves on an arbitrary small site is a topos. (A more detailed proof can be extracted from the results of Sections A4.3 and A4.4, in particular A4.3.5 and A4.4.5.) For technical reasons which will become apparent shortly, we now wish to prove a more general version of this result, in which the hypothesis of smallness is relaxed slightly. In order to do this, we shall first need to establish a result (the Comparison Lemma) of considerable importance in its own right, which provides a criterion for determining when two sites give rise to equivalent categories of

sheaves. We shall assume in this section that all coverages are Grothendieck coverages, unless stated otherwise.

Definition 2.2.1 Let $(\mathcal{C}, J)$ be a site. We say a subcategory $\mathcal{D}$ of $\mathcal{C}$ is *J-dense* (or simply *dense*, if J is obvious from the context) if

(i) every object U of $\mathcal{C}$ has a covering sieve $R \in J(U)$ generated by morphisms whose domains are in $\mathcal{D}$ (equivalently, the sieve generated by all morphisms $V \to U$ with $V \in \text{ob } \mathcal{D}$ is J-covering); and

(ii) for any morphism $f: U \to V$ in $\mathcal{C}$ with $V \in \text{ob } \mathcal{D}$, there is a covering sieve $R \in J(U)$ generated by morphisms $g: W \to U$ for which the composite fg is in $\mathcal{D}$ (equivalently, the family of all such morphisms generates a J-covering sieve).

It is easy to see that, in the presence of condition (i) of 2.2.1, condition (ii) may be reduced to the special case in which the domain U as well as the codomain V of f is in $\mathcal{D}$; hence if $\mathcal{D}$ is a *full* subcategory condition (ii) may be omitted altogether. (On the other hand, given condition (ii) as we have stated it, condition (i) becomes redundant provided every object of $\mathcal{C}$ admits a morphism to some object of $\mathcal{D}$ – for example if $\mathcal{D}$ contains a terminal object of $\mathcal{C}$.) In practice, the Comparison Lemma is most often used for full subcategories, and many texts only define denseness in this case; however, the extra generality afforded by the definition we have given is occasionally useful – we shall see an instance of its use in 5.2.5 below.

Given any subcategory $\mathcal{D}$ of $\mathcal{C}$ and a coverage J on $\mathcal{C}$, we define the *induced coverage* $J_{\mathcal{D}}$ on $\mathcal{D}$ by setting $J_{\mathcal{D}}(V)$ to be the collection of all sieves $R \cap \text{mor } \mathcal{D}$, where $R \in J(V)$. It is straightforward to verify that $J_{\mathcal{D}}$ satisfies the closure conditions (C), (L) and (M) if J does.

Lemma 2.2.2 *Let $(\mathcal{C}, J)$ be a site, and let $\mathcal{D}$ be a subcategory of $\mathcal{C}$ satisfying condition* (ii) *of* 2.2.1. *Then*

(i) *A sieve S on an object V of $\mathcal{D}$ is $J_{\mathcal{D}}$-covering iff the sieve $\overline{S}$ in $\mathcal{C}$ generated by the members of S is J-covering.*

(ii) *For any J-sheaf A on $\mathcal{C}$, the restriction of A to $\mathcal{D}$ is a $J_{\mathcal{D}}$-sheaf.*

Proof (i) The right-to-left implication is clear, since $\overline{S} \cap \text{mor } \mathcal{D} = S$. For the converse, suppose that $S = R \cap \text{mor } \mathcal{D}$ for some J-covering sieve R. Then, for any $f: U \to V \in R$, $f^*(\overline{S})$ contains all morphisms $g: W \to U$ with $fg \in \text{mor } \mathcal{D}$, and is therefore J-covering; so $\overline{S}$ is J-covering by (L).

(ii) Suppose $S \in J_{\mathcal{D}}(V)$, and let $(s_f \in A(\text{dom } f) \mid f \in S)$ be a compatible family of elements for S. By (i), it is sufficient to show that $(s_f \mid f \in S)$ extends

to a compatible family for $\overline{S}$, i.e. that if we have a commutative diagram in $\mathcal{C}$

$$\begin{tikzcd} W \arrow[r, "g_1"] \arrow[d, "g_2"] & U_1 \arrow[d, "f_1"] \\ U_2 \arrow[r, "f_2"] & V \end{tikzcd}$$

with $f_1, f_2 \in S$ then $A(g_1)(s_{f_1}) = A(g_2)(s_{f_2})$. But W can be J-covered by morphisms $h \colon X \to W$ such that the composites $g_1 h$ and $g_2 h$ both lie in $\mathcal{D}$, so that the images of the two elements above under $A(h)$ both coincide with $s_{f_1 g_1 h}$. Since A is a J-sheaf, it follows that the two elements are equal. □

Thus, under the hypotheses of Lemma 2.2.2, the functor $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ which sends A to its restriction to $\mathcal{D}$ itself restricts to a functor $\mathbf{Sh}(\mathcal{C}, J) \to \mathbf{Sh}(\mathcal{D}, J_{\mathcal{D}})$. To get a functor in the opposite direction, we shall need to introduce some smallness hypotheses.

Theorem 2.2.3 (Comparison Lemma) *Let $(\mathcal{C}, J)$ be a site such that $\mathcal{C}$ is locally small, and let $\mathcal{D}$ be a small dense subcategory of $\mathcal{C}$. Then the restriction functor $\mathbf{Sh}(\mathcal{C}, J) \to \mathbf{Sh}(\mathcal{D}, J_{\mathcal{D}})$ just defined is (one half of) an equivalence of categories.*

Proof Let $B \colon \mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ be a functor. For each object U of $\mathcal{C}$, let $(\mathcal{D} \downarrow U)$ denote the category whose objects are morphisms $f : V \to U$ in $\mathcal{C}$ with $V \in \mathrm{ob}\, \mathcal{D}$, and whose morphisms $f \to f'$ are morphisms $g \colon \mathrm{dom}\, f \to \mathrm{dom}\, f'$ in $\mathcal{D}$ such that $f'g = f$. The hypotheses on $\mathcal{C}$ and $\mathcal{D}$ ensure that $(\mathcal{D} \downarrow U)$ is a small category, and hence we can form the limit $A(U)$ of the composite

$$(\mathcal{D} \downarrow U)^{\mathrm{op}} \longrightarrow \mathcal{D}^{\mathrm{op}} \xrightarrow{\;B\;} \mathbf{Set},$$

where the first factor is the forgetful functor $f \mapsto \mathrm{dom}\, f$. The assignment $U \mapsto A(U)$ becomes a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ in an obvious way: a morphism $U \to U'$ induces a functor $(\mathcal{D} \downarrow U) \to (\mathcal{D} \downarrow U')$ over $\mathcal{D}$, and hence a comparison map between the two limits. Similarly, the assignment $B \mapsto A$ itself becomes a functor $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. (Of course, this is just the usual construction of the right Kan extension of B along the inclusion functor $\mathcal{D}^{\mathrm{op}} \to \mathcal{C}^{\mathrm{op}}$.)

Next, we note that the denseness of $\mathcal{D}$ ensures that any sheaf A on $\mathcal{C}$ is isomorphic to the Kan extension of its restriction to $\mathcal{D}$. For if we are given an element of the above limit (with B taken to be the restriction of A), i.e. a family of elements $(s_f \in B(\mathrm{dom}\, f) \mid f \in \mathrm{ob}\, (\mathcal{D} \downarrow U))$ satisfying the appropriate compatibility conditions, then an argument like that in the proof of 2.2.2(ii) ensures that this family extends to a compatible family for the sieve (in $\mathcal{C}$) generated by all morphisms from objects of $\mathcal{D}$ to U, and hence that there is a unique $s \in A(U)$ with $A(f)(s) = s_f$ for all f.

Similarly, any $J_{\mathcal{D}}$-sheaf B on $\mathcal{D}$ is isomorphic to the restriction of its right Kan extension. If $\mathcal{D}$ is a full subcategory of $\mathcal{C}$ this is trivial, since 1_V is a terminal object of $(\mathcal{D}\downarrow V)$ for every $V \in \text{ob}\,\mathcal{D}$. In general, if we are given an element $(s_f \in B(W) \mid (f\colon W \to V) \in \text{ob}\,(\mathcal{D}\downarrow V))$ of $A(V)$ where $V \in \text{ob}\,\mathcal{D}$, then for each f the morphisms $g\colon X \to W$ in $\mathcal{D}$ for which $fg \in \text{mor}\,\mathcal{D}$ generate a $J_{\mathcal{D}}$-covering sieve, so that s_f is uniquely determined by the s_{fg} for all such g – and these are in turn determined by s_{1_V}. So we have a canonical bijection from $B(V)$ to $A(V)$.

Thus, in view of Lemma 2.2.2, it remains only to prove that the right Kan extension of a $J_{\mathcal{D}}$-sheaf on $\mathcal{D}$ is indeed a J-sheaf on $\mathcal{C}$. For this, we can conveniently use the characterization of sheaves given by 2.1.4: suppose A is the right Kan extension of B, and we are given $R \in J(U)$ and a natural transformation $\alpha\colon R \to A$. We have to produce an element of $A(U)$; but by definition this consists of a compatible family of elements $(x_f \in B(V) \mid f\colon V \to U, V \in \text{ob}\,\mathcal{D})$. For each such f, the sieve $f^*(R) \cap \text{mor}\,\mathcal{D}$ is $J_{\mathcal{D}}$-covering on V, and α defines (by restriction) a compatible family of elements of B relative to this cover; so we use the $J_{\mathcal{D}}$-sheaf axiom for B to construct the required element x_f. The rest is straightforward verification. □

We shall say a site $(\mathcal{C}, J)$ is *essentially small* if $\mathcal{C}$ is locally small and has a small J-dense subcategory. We note that, although we required the subcategory $\mathcal{D}$ to be small in the statement of 2.2.3, the result has an immediate extension to the case where $\mathcal{D}$ (with the induced coverage $J_{\mathcal{D}}$) is an essentially small site: for if $\mathcal{D}'$ is a small $J_{\mathcal{D}}$-dense subcategory of such a $\mathcal{D}$, it is easy to see that $\mathcal{D}'$ is also J-dense in $\mathcal{C}$, and hence both $\mathbf{Sh}(\mathcal{C})$ and $\mathbf{Sh}(\mathcal{D})$ are equivalent to $\mathbf{Sh}(\mathcal{D}')$.

It is also possible to formulate a still more general version of the Comparison Lemma, for functors $F\colon \mathcal{D} \to \mathcal{C}$ between the underlying categories of sites which are not even faithful; but we shall not need this extra generality.

Examples 2.2.4 (*a*) Let X be a locale, and consider the category $\mathbf{LH}/X$, equipped with its canonical coverage J. By the proof of 1.3.11, any object of $\mathbf{LH}/X$ is a colimit of a diagram whose vertices are open inclusions $U \rightarrowtail X$, and these colimits are stable under pullback; so the full subcategory (isomorphic to $\mathcal{O}(X)$) whose objects are these inclusions is J-dense in $\mathbf{LH}/X$. Moreover, the induced coverage $J_{\mathcal{O}(X)}$ on $\mathcal{O}(X)$ is readily seen to be the (canonical) coverage used in the original definition of a sheaf on a locale; so we deduce that $\mathbf{Sh}(\mathbf{LH}/X, J)$ is equivalent to $\mathbf{Sh}(X)$ (and hence, by 1.3.11, to $\mathbf{LH}/X$ itself). This example is a special case of a general phenomenon which we shall observe in 2.2.7 below.

(*b*) More generally, by a *basis* for a locale X we mean a subset $\mathcal{B}$ of $\mathcal{O}(X)$ which generates it as a complete join-semilattice, i.e. such that every element of $\mathcal{O}(X)$ is expressible as a join of members of $\mathcal{B}$. Then the inclusion $\mathcal{B} \rightarrowtail \mathcal{O}(X)$ satisfies the conditions of 2.2.1 with respect to the canonical coverage J on $\mathcal{O}(X)$, so we may also regard $\mathbf{Sh}(X)$, up to equivalence, as $\mathbf{Sh}(\mathcal{B}, J_{\mathcal{B}})$. For example, if $\mathcal{O}(X) = IA$ is the frame of ideals of a distributive lattice A (cf. 1.1.16(*e*)), and $\mathcal{B}$ is the subset of principal ideals (isomorphic as a poset to A), then the induced

coverage $J_{\mathcal{B}}$ is simply the coherent coverage, since any covering of a principal ideal has a finite subcover. Thus we obtain the result that *the topos of sheaves on a distributive lattice A for its coherent coverage is equivalent to the topos of sheaves on the locale corresponding to the frame of ideals of A.*

(c) Let G be a topological group, and consider the topos $\mathbf{Cont}(G)$ of continuous G-sets, defined in A2.1.6, as a site equipped with its canonical coverage C. It is easy to see that the full subcategory consisting of the transitive G-sets is dense, since any continuous G-set is the disjoint union of its G-orbits; so we may make the transitive continuous G-sets into the objects of a site $(\mathcal{C}, J)$ such that $\mathbf{Sh}(\mathcal{C}, J) \simeq \mathbf{Sh}(\mathbf{Cont}(G), C)$. (As with example ($a$) above, we shall see in 2.2.7 below that the latter is equivalent to $\mathbf{Cont}(G)$ itself.)

Clearly, any transitive continuous G-set is isomorphic to the set $G : H$ of left cosets of some open subgroup H of G, with G acting by left translation; so we may take these G-sets to be the objects of our category $\mathcal{C}$. Alternatively, we can take the objects to be the open subgroups themselves, with morphisms $H \to K$ labelled by left cosets gK of K such that $H \subseteq gKg^{-1}$ – note that this makes sense, since the conjugate gKg^{-1} depends only on the coset gK, and not on the choice of representative g. (The idea is that $gK \colon H \to K$ corresponds to the unique G-equivariant map $G : H \to G : K$ sending H to the coset gK.) The composite of

$$H \xrightarrow{gK} K \xrightarrow{hL} L$$

is ghL; again, we need to check that this is well-defined, but the inclusion $K \subseteq hLh^{-1}$, or equivalently $h^{-1}Kh \subseteq L$, implies that $KhL = hL$, and hence $ghL = gKhL$ depends only on gK and hL. And since every morphism between transitive G-sets is epic in $\mathbf{Cont}(G)$, the coverage J on this category has the property that every inhabited sieve is covering (cf. 2.1.12(c)). We leave to the reader the task of verifying that $\mathcal{C}$ satisfies the right Ore condition.

As in (b), we may replace the set of all open subgroups in this example by a basis for the open subgroups, i.e. a family $\mathcal{B}$ of open subgroups such that every open subgroup contains a member of $\mathcal{B}$. And in 5.3.9 below we shall see that this site can still be used in the case of a non-spatial localic group G (that is, a group object in $\mathbf{Loc}$); we already knew that the continuous actions of such a group formed a topos, by B3.4.14(b), but we did not previously have a description of it as a category of sheaves on a site. (Nontrivial examples of localic groups with no points other then the identity element do exist: for a simple one, revisit the non-spatial locale constructed as an inverse limit in 1.2.9, and replace the sets X_i appearing in the inverse system by the free groups which they generate.)

(d) Let $\mathcal{C}$ be any locally small category, $\mathcal{D}$ a small full subcategory of $\mathcal{C}$, and define a coverage J on $\mathcal{C}$ by saying that a sieve R on U is J-covering iff it contains all the morphisms from objects of $\mathcal{D}$ to U. It is easily verified that this is indeed a (Grothendieck) coverage, that $\mathcal{D}$ is a J-dense subcategory, and that the induced coverage $J_{\mathcal{D}}$ is the trivial one whose only covers are the maximal sieves M_V, $V \in \mathrm{ob}\, \mathcal{D}$. Thus 2.2.3 yields an equivalence $\mathbf{Sh}(\mathcal{C}, J) \simeq [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$.

In the case when $\mathcal{C}$ is also small, this equivalence identifies the geometric inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ with that induced as in A4.2.12(*b*) by the full embedding $\mathcal{D}^{\mathrm{op}} \to \mathcal{C}^{\mathrm{op}}$. (We shall return to this example in 2.2.19 below.)

(*e*) Let $(\mathcal{C}, J)$ be a site, and let $\mathcal{C}'$ be the full subcategory on those objects of $\mathcal{C}$ which are *not* J-covered by the empty sieve. Then it is clear that $\mathcal{C}'$ is J-dense in $\mathcal{C}$. Moreover, it follows easily from (L) that the induced coverage J' on $\mathcal{C}'$ has all its sieves nonempty; so (provided we assume classical logic in our base topos **Set** – a point to which we shall return in Section C3.1 below) we see that *every category of sheaves on a site is equivalent to one on a site where every covering sieve is nonempty.* This condition on a site $(\mathcal{C}, J)$ correponds to saying that $\mathbf{Sh}(\mathcal{C}, J)$ is a dense subtopos of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, i.e. that the initial object of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ (the constant functor with value 0) is a sheaf; so we can regard the italicized statement above as an analogue for toposes of the fact, noted in 1.2.7, that any locale is a dense sublocale of a space.

Remark 2.2.5 We may now prove, as promised in the last section, that *if J and K are two Grothendieck coverages on a category $\mathcal{C}$ such that the sites $(\mathcal{C}, J)$ and $(\mathcal{C}, K)$ are both essentially small, and* $\mathbf{Sh}(\mathcal{C}, J)$ *coincides with* $\mathbf{Sh}(\mathcal{C}, K)$ *as a subcategory of* $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, *then* $J = K$. For we may find a small (full) subcategory $\mathcal{D}$ of $\mathcal{C}$ which is dense for both J and K; then $\mathbf{Sh}(\mathcal{D}, J_{\mathcal{D}})$ and $\mathbf{Sh}(\mathcal{D}, K_{\mathcal{D}})$ coincide as subcategories of $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$, and so $J_{\mathcal{D}} = K_{\mathcal{D}}$ by 2.1.11. But, given that J satisfies (L), it is easily seen that we may recover it from $J_{\mathcal{D}}$, as follows: a sieve R on an object U of $\mathcal{C}$ is J-covering iff, for each $f: V \to U$ with $V \in \mathrm{ob}\, \mathcal{D}$, the sieve $f^*R \cap \mathrm{mor}\, \mathcal{D}$ is $J_{\mathcal{D}}$-covering. Hence $J_{\mathcal{D}} = K_{\mathcal{D}}$ implies $J = K$.

We next recall

Proposition 2.2.6 *For any small site $(\mathcal{C}, J)$,* $\mathbf{Sh}(\mathcal{C}, J)$ *is a topos, and has all small limits and colimits. Moreover, it is reflective in* $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

Proof As we remarked earlier in this section, we have already proved this in Part A, but we shall repeat the proof here, expanding on the brief outline given in A2.1.10.

The first step, the construction of small limits in $\mathbf{Sh}(\mathcal{C}, J)$, does not in fact require any smallness assumption; it is immediate from the definition that the class of functors which satisfy the sheaf axiom for a given family is closed under all (pointwise) limits in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, and hence that it has all small limits.

From now on, we assume that $\mathcal{C}$ is small. In A1.5.5, we saw that $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is cartesian closed; we shall show that $\mathbf{Sh}(\mathcal{C}, J)$ is an exponential ideal in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, and hence cartesian closed in its own right. Suppose given a J-sheaf A, an arbitrary functor $B: \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, and a natural transformation $\theta: R \to A^B$ where $R \in J(U)$ for some $U \in \mathrm{ob}\, \mathcal{C}$. Then extending θ to a natural transformation $\phi: \mathcal{C}(-, U) \to A^B$ is equivalent to extending $\overline{\theta}: R \times B \to A$ to a natural transformation $\overline{\phi}: \mathcal{C}(-, U) \times B \to A$; we shall show that the latter is uniquely possible. Given any $f: V \to U$ and any $x \in B(V)$, the sieve f^*R is J-covering, and the pairs $((fg, B(g)(x)) \mid g \in f^*R)$ form a compatible family of elements of

$R \times B$, so their images under $\overline{\theta}$ form a compatible family of elements of A. We define $\overline{\phi}_V(f, x)$ to be the element of $A(V)$ obtained by applying the sheaf axiom to this family; it is easy to check that $\overline{\phi}$ is natural, and that it is the unique natural transformation extending $\overline{\theta}$.

We saw in A1.6.6 that the subobject classifier Ω in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is given by $\Omega(U) = \{\text{sieves on } U\}$ with $\Omega(f)(R) = f^*(R)$. We define a sieve $R \in \Omega(U)$ to be *J-closed* if every $f \colon V \to U$ such that $f^*(R) \in J(V)$ belongs to R, i.e. if $f^*(R) \in J(V)$ implies $f^*(R) = M_V$. It is clear from the definition that the assignment $U \mapsto \{J\text{-closed sieves on } U\}$ defines a subfunctor Ω_J of Ω; moreover, if A is a sheaf, then a subfunctor A' of A is a sheaf iff, for any $x \in A(U)$, the sieve $\{f \colon V \to U \mid A(f)(x) \in A'(V)\}$ is J-closed, i.e. iff the classifying map of $A' \rightarrowtail A$ factors through $\Omega_J \rightarrowtail \Omega$. So, provided we can show that Ω_J itself is a J-sheaf, it will serve as a subobject classifier for $\mathbf{Sh}(\mathcal{C}, J)$. But, if $R \in J(U)$ and we are given a compatible family $(S_f \mid f \in R)$ of J-closed sieves on the domains of the members of R, then there is a unique closed sieve S on U with $f^*(S) = S_f$ for all f, namely the sieve of all $g \colon V \to U$ such that $S_{gf} = M_W$ for all $f \colon W \to V \in g^*(R)$.

The existence of all small colimits in $\mathbf{Sh}(\mathcal{C}, J)$ may be shown either by appealing to the Monadicity Theorem (A2.2.7) or by showing that $\mathbf{Sh}(\mathcal{C}, J)$ is reflective in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. We did the latter in A4.4.4, but we shall give a different construction of the reflection here. Given a functor $A \colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, we define $A^+(U)$ to be the colimit, over the ordered set $J(U)$ of J-covering sieves on U, of the sets of morphisms $R \to A$; equivalently, an element of $A^+(U)$ is an equivalence class of pairs (R, s) where $R \in J(U)$ and $s = (s_f \in A(\mathrm{dom}\ f) \mid f \in R)$ is a compatible family of elements of A relative to R, the equivalence relation identifying (R, s) and (R', s') iff there is a J-covering sieve $R'' \subseteq R \cap R'$ on which the restrictions of s and s' agree. Given $f \colon V \to U$ in $\mathcal{C}$, it is easy to see that we obtain a well-defined mapping $A^+(U) \to A^+(V)$ sending the equivalence class of (R, s) to that of $(f^*(R), (s_{fg} \mid g \in f^*(R)))$; thus A^+ becomes a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$. Moreover, the assignment $A \mapsto A^+$ is itself functorial in an obvious way, and we have a natural transformation $\eta_A \colon A \to A^+$ sending $s \in A(U)$ to the equivalence class of $(M_U, (A(f)(s) \mid f \in M_U))$, which is an isomorphism iff A is a J-sheaf. Hence every morphism from a general A to a sheaf factors uniquely through η_A.

For a general A, A^+ need not be a sheaf, but it is at least J-separated. To see this, let (R, s) and (R', s') represent two elements of $A^+(U)$ which agree when restricted to the domain of every morphism in a third J-covering sieve R'' on U. Thus, for every $f \colon V \to U$ in R'', we have a J-covering sieve S_f on V, contained in $f^*(R) \cap f^*(R')$, such that $s_{fg} = s'_{fg}$ for every $g \in S_f$. But then, by (L), the sieve consisting of all the composites fg, as f ranges over R'' and g over S_f, is J-covering on U; so (R, s) and (R', s') represent the same element of $A^+(U)$.

To complete the construction of the left adjoint, it suffices to show that if A is separated then A^+ is a sheaf; for then we shall know that A^{++} is a sheaf for any functor A, and every morphism from A to a sheaf factors uniquely through the composite $\eta_{A^+} \circ \eta_A \colon A \to A^+ \to A^{++}$. So suppose A is separated; then we

see that if (R, s) and (R', s') represent the same element of $A^+(U)$, then s and s' must agree on the whole of $R \cap R'$ (rather than on some covering sieve contained in $R \cap R'$), since for any $f \in R \cap R'$ the elements s_f and s'_f agree on a J-covering sieve. Thus each element of $A^+(U)$ has a canonical representative, namely the union of all the families which represent it. So if we now have a covering sieve R on U and a compatible family of elements $(t_f \in A^+(\mathrm{dom}\ f) \mid f \in R)$, the canonical representatives (S_f, s_f) of the t_f can be combined into a single compatible family of elements of A indexed by the sieve $\{fg \mid f \in R, g \in S_f\}$, which is again J-covering by (L). In this way we obtain an element t of $A^+(U)$ such that $A^+(f)(t) = t_f$ for each $f \in R$; so A^+ satisfies the 'existence' as well as the 'uniqueness' part of the definition of a sheaf. □

We recall that in Section A4.4 we not only constructed a left adjoint for the inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, but also proved that it is cartesian (i.e. preserves finite limits). In terms of the construction given in the proof of 2.2.6, the same result may be obtained by observing that the poset $J(U)$ of J-covering sieves on U is codirected, since it has finite intersections; so the fact that filtered colimits and finite limits commute in **Set** implies that the functor $A \mapsto A^+$ is cartesian. As we remarked at the end of Section A4.4, there is also an 'internal' version of this construction of the associated sheaf functor, applicable to any local operator in an elementary topos; cf. [501].

We note that a topos of the form $\mathbf{Sh}(\mathcal{C}, J)$, as well as having all small colimits, is locally small; hence by A4.1.9 it admits a geometric morphism to **Set**. It is also well-powered (since subobjects of an object A correspond bijectively to morphisms $A \to \Omega$) and hence an ∞-pretopos as defined in Section A1.4. In the converse direction, we have

Proposition 2.2.7 *Let $\mathcal{C}$ be an ∞-pretopos with a separating set. Then the canonical coverage J makes $(\mathcal{C}, J)$ into an essentially small site, and every J-sheaf on $\mathcal{C}$ is representable.*

Proof By 2.1.12(e), we know that the canonical coverage consists of all sieves which contain small epimorphic families. But the definition of separating set says that, for any $U \in \mathrm{ob}\ \mathcal{C}$, the family of all morphisms from objects in the separating set to U is (small and) epimorphic, so the full subcategory $\mathcal{D}$ on these objects is J-dense. Thus $(\mathcal{C}, J)$ is essentially small.

Let $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ be the factorization through $\mathbf{Sh}(\mathcal{C}, J)$ of the Yoneda embedding $U \mapsto \mathcal{C}(-, U)$. By the argument of A2.1.11(a) and (b), l preserves covers and small unions, i.e. it is a geometric functor. We must show that every object of $\mathbf{Sh}(\mathcal{C}, J)$ is isomorphic to one in the image of l. Given a sheaf A, let S be the set of pairs (V, x) with $V \in \mathrm{ob}\ \mathcal{D}$ and $x \in A(V)$, and let W be the coproduct (in $\mathcal{C}$) of the S-indexed family $(V \mid (V, x) \in S)$. From the proof of 2.2.3, we know that any J-sheaf B is isomorphic to the right Kan extension of its restriction to $\mathcal{D}$, and so an element $y \in B(U)$ (for any $U \in \mathrm{ob}\ \mathcal{C}$) is uniquely determined by the family of elements $(B(f)(y) \mid f\colon V \to U, V \in \mathrm{ob}\ \mathcal{D})$. It follows

that two natural transformations $A \rightrightarrows B$ which agree at all objects of $\mathcal{D}$ must be equal, and hence the natural transformation $l(W) \cong \coprod_{(V,x)\in S} l(V) \to A$ whose (V, x)th component is (induced by) x is epimorphic in $\mathbf{Sh}(\mathcal{C}, J)$.

Now form the kernel-pair $R \rightarrowtail l(W) \times l(W)$ of $l(W) \twoheadrightarrow A$. By the same argument, we can find an epimorphism $l(X) \twoheadrightarrow R$. But l (is full and) preserves images, so we have $R \cong l(I)$, where I is the image of the corresponding pair $X \to W \times W$ in $\mathcal{C}$. And since epimorphisms in $\mathbf{Sh}(\mathcal{C}, J)$ are regular, and l preserves coequalizers of equivalence relations, we now have $A \cong l(Q)$, where Q is the coequalizer of $I \rightrightarrows W$ in $\mathcal{C}$. So the result is proved. □

Summarizing the results of this section so far, we now have the characterization of sheaf toposes known as Giraud's Theorem [405].

Theorem 2.2.8 *For a category $\mathcal{E}$, the following are equivalent:*

(i) *There exists an essentially small site $(\mathcal{C}, J)$ such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$.*

(ii) *There exists a small site $(\mathcal{C}, J)$ such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$.*

(iii) *There exists a standard site $(\mathcal{C}, J)$ such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$.*

(iv) *$\mathcal{E}$ is a cocomplete, locally small topos with a separating set of objects.*

(v) *$\mathcal{E}$ is a topos, and there exists a bounded geometric morphism $\mathcal{E} \to \mathbf{Set}$ (cf.* B3.1.7*).*

(vi) *$\mathcal{E} \simeq \mathbf{Map}(\mathcal{A})$, where $\mathcal{A}$ is a tabular effective positive geometric allegory with a strong separator (cf.* A3.4.11*).*

(vii) *$\mathcal{E}$ is an ∞-pretopos with a separating set of objects.*

(viii) *$\mathcal{E}$ is locally small and has a small subcategory which is dense for the canonical coverage, and every sheaf for the canonical coverage is representable.*

(ix) *$\mathcal{E}$ is locally presentable (cf.* D2.3.4*), locally cartesian closed and balanced, and is effective as a regular category.*

Proof (i) ⇒ (ii) follows from 2.2.3, and (iii) ⇒ (ii) is obvious. (ii) ⇒ (iv) was all proved in 2.2.6, except for the separating set of objects; but it is well known that the hom-functors $\mathcal{C}(-, U), U \in \text{ob}\,\mathcal{C}$, form a separating set for $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$ if $\mathcal{C}$ is small, and since $\mathbf{Sh}(\mathcal{C}, J)$ is reflective in $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$ their images under the reflector form a separating set for $\mathbf{Sh}(\mathcal{C}, J)$. Now (iii) ⇒ (ix) because D2.3.7 shows that any category satisfying (iii) is locally presentable, and the other hypotheses of (ix) are satisfied by any topos. The equivalence of (iv) and (v) was established in B3.1.8(*b*) (cf. also A4.1.9), and that of (v) and (vi) in B3.1.9. (iv) ⇒ (vii) because a locally small topos is well-powered, and a cocomplete well-powered topos is an ∞-pretopos. Similarly, (ix) ⇒ (vii) because the infinitary analogue of A1.5.13 shows that $\mathcal{E}$ is a geometric category (it is well-powered by A1.4.17) and hence an ∞-pretopos by the infinitary analogue of A1.5.14, and it has a generating set by the definition of local presentability. (vii) ⇒ (viii) by 2.2.7 and the fact (which follows from 2.1.12(*e*)) that any separating set for an ∞-pretopos defines

a dense full subcategory for the canonical coverage. Assuming (viii), let J be the canonical coverage on $\mathcal{E}$; then, since the Yoneda embedding $l: \mathcal{E} \to \mathbf{Sh}(\mathcal{E}, J)$ is full and faithful, (viii) says that l is a weak equivalence, and so (with choice) we have an equivalence $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{E}, J)$. Hence (i) holds.

It remains to show that (iii) is implied by the other conditions. First, we note that if $\mathcal{D}$ is a dense subcategory then so is the full subcategory $\overline{\mathcal{D}}$ of $\mathcal{E}$ on the same objects; and the induced coverage $J_{\overline{\mathcal{D}}}$ on the latter will be subcanonical, since all the functors $\overline{\mathcal{D}}(-, V)$ ($V \in \text{ob } \mathcal{D}$) are sheaves for it. Moreover, $\overline{\mathcal{D}}$ is small if $\mathcal{D}$ is, since $\mathcal{E}$ is locally small; and we may further close it under finite limits in $\mathcal{E}$ without destroying its smallness (and clearly without destroying its denseness either). □

Definition 2.2.9 A category satisfying the equivalent conditions of 2.2.8 is called a *Grothendieck topos.*

Remark 2.2.10 We have seen that a Grothendieck topos is complete, cocomplete, locally small and well-powered, and it has a separating set. It is also well-copowered (because isomorphism classes of epimorphisms with domain A correspond bijectively to equivalence relations on A), and has a coseparating set (in fact a single coseparator), by B3.1.13. Thus we are always free to appeal to the Special Adjoint Functor Theorem (B2.4.6) when considering functors between Grothendieck toposes: that is, *a functor between Grothendieck toposes has a left (resp. right) adjoint iff it preserves all small limits (resp. colimits).* In particular, we note that if $F: \mathcal{E} \to \mathcal{F}$ is a geometric functor between Grothendieck toposes, then it preserves all small colimits (by the argument of A1.4.19), and so has a right adjoint; thus it is the inverse image of a geometric morphism $\mathcal{F} \to \mathcal{E}$. (This is essentially a restatement of something we observed in A4.1.18.)

Another consequence of the 'smallness condition' built into the definition of a Grothendieck topos is also worth stating explicitly. Although the hom-categories of the 2-category $\mathfrak{BTop}/\mathbf{Set}$ of Grothendieck toposes need not be small, they do have an important 'smallness property':

Lemma 2.2.11 *Let $\mathcal{E}$ and $\mathcal{F}$ be Grothendieck toposes. Then there exists a set K of geometric morphisms $\mathcal{F} \to \mathcal{E}$ which is 'relatively conservative': that is, if $u: A' \rightarrowtail A$ is a monomorphism in $\mathcal{E}$ such that $f^*(u)$ is an isomorphism for all $f \in K$, then $g^*(u)$ is an isomorphism for* all *geometric morphisms $g: \mathcal{F} \to \mathcal{E}$.*

Proof Let $(\mathcal{C}, J)$ be a small site such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$, and as usual let $l: \mathcal{C} \to \mathcal{E}$ denote the composite of the Yoneda embedding and the associated sheaf functor. For each $U \in \text{ob } \mathcal{C}$ and each subobject $m: R \rightarrowtail l(U)$ in $\mathcal{E}$, if there exists a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ such that $f^*(m)$ is not an isomorphism, then choose one and call it f_m; let K be the set of all such f_m, as m ranges over (a representative set of) subobjects of objects of the form $l(U)$. Now let $u: A' \rightarrowtail A$ be an arbitrary monomorphism in $\mathcal{E}$, and suppose $g^*(u)$ is not an isomorphism for some $g: \mathcal{F} \to \mathcal{E}$. Since A can be expressed as a colimit of objects

of the form $l(U)$, and since g^* preserves colimits and pullbacks, there must exist $x\colon l(U) \to A$ in $\mathcal{E}$ for which the pullback

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail, "m"] & A' \arrow[d, tail, "u"] \\ l(U) \arrow[r, "x"] & A \end{tikzcd}$$

is such that $g^*(m)$ is not an isomorphism. But then we have $f_m \in K$ such that $f_m^*(m)$ is not an isomorphism, and hence $f_m^*(u)$ is not an isomorphism either. □

Lemma 2.2.11 is most often of use in the particular case $\mathcal{F} = \mathbf{Set}$. By analogy with the case of sheaves on a locale, we define a *point* of a Grothendieck topos $\mathcal{E}$ to be a geometric morphism $\mathbf{Set} \to \mathcal{E}$, and we say that $\mathcal{E}$ *has enough points* if the class of all inverse image functors $\mathcal{E} \to \mathbf{Set}$ is jointly conservative. We may thus deduce

Corollary 2.2.12 *A Grothendieck topos $\mathcal{E}$ has enough points iff there exists a surjection $\mathbf{Set}/K \to \mathcal{E}$ for some set K.*

Proof By B3.4.1 and the remarks following it (plus A1.1.6, if you insist), specifying a geometric morphism $\mathbf{Set}/K \to \mathcal{E}$ is equivalent to specifying a K-indexed family of points of $\mathcal{E}$. Moreover, the former is surjective iff the inverse images of the latter are jointly conservative. So the result is immediate from 2.2.11. □

In D2.3.12 we prove a stronger version of 2.2.11: namely, that $\mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E})$ contains a set of objects K such that every object is expressible as a filtered colimit of members of K. (We recall from B3.4.8 – though it also follows easily from 2.2.10 – that filtered colimits in $\mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E})$ are computed 'pointwise' at the level of inverse image functors, so 2.2.11 does indeed follow from this assertion.)

Next, we digress to consider the analogue for quasitoposes of Giraud's Theorem. By a *bisite*, we mean a category $\mathcal{C}$ equipped with two coverages J and K, such that $J \subseteq K$. We say a functor $A\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is *biseparated* for the pair (J, K) if it satisfies the sheaf axiom for all covers in J and is separated for all covers in K; and we write $\mathbf{Bisep}(\mathcal{C}, J, K)$ for the category of all such functors (and natural transformations between them). If $\mathcal{C}$ is small, it is easy to see that these functors are exactly the objects of $\mathbf{Sh}(\mathcal{C}, J)$ which are separated (in the sense of A4.3.4) for the local operator k corresponding to the subtopos $\mathbf{Sh}(\mathcal{C}, K) \subseteq \mathbf{Sh}(\mathcal{C}, J)$; hence by A4.4.5 $\mathbf{Bisep}(\mathcal{C}, J, K)$ is a quasitopos. Moreover, it is reflective in $\mathbf{Sh}(\mathcal{C}, J)$ and hence in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; so it is complete and cocomplete, and has a generating set consisting of the reflections of the representable functors.

One more definition: we say an equivalence relation $R \rightrightarrows A$ in a regular category is *strong* if the corresponding monomorphism $R \rightarrowtail A \times A$ is a cocover (i.e. a strong monomorphism). Clearly, every effective equivalence relation is strong, since the indicated monomorphism is an equalizer. We say that a regular category is *quasi-effective* if every strong equivalence relation is effective. The proof given in A2.4.1 that a topos is effective as a regular category may easily be adapted to prove that a quasitopos is quasi-effective.

Theorem 2.2.13 *For a category $\mathcal{E}$, the following conditions are equivalent:*

(i) *There exists a Grothendieck topos $\mathcal{F}$ and a local operator k on $\mathcal{F}$ such that $\mathcal{E} \simeq \mathbf{sep}_k(\mathcal{F})$.*

(ii) *$\mathcal{E}$ is equivalent to $\mathbf{Bisep}(\mathcal{C}, J, K)$ for some small bisite $(\mathcal{C}, J, K)$ (whose underlying category may be taken to have finite limits).*

(iii) *$\mathcal{E}$ is a locally small, cocomplete quasitopos with a generating set.*

(iv) *$\mathcal{E}$ is locally presentable, locally cartesian closed and quasi-effective.*

Proof The proof of (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) is mostly contained in the discussion before the statement of the theorem; the only thing missing is the local presentability, which may be established as in D2.3.7. So it remains to prove (iv) $\Rightarrow$ (ii).

Suppose $\mathcal{E}$ is locally κ-presentable. Then by D2.3.4 we can represent it as the category of κ-continuous functors $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, where $\mathcal{C}$ is a small category (equivalent to the full subcategory of κ-presentable objects of $\mathcal{E}$) with κ-small colimits. We shall take this $\mathcal{C}$ as the underlying category of our bisite, and regard it as a full subcategory of $\mathcal{E}$. The coverages J and K are defined as those induced on $\mathcal{C}$ by the coverages on $\mathcal{E}$ consisting of sieves containing small effective-epimorphic families, and small epimorphic families, respectively: note that, since $\mathcal{E}$ is a geometric category by A1.5.13, these families are indeed stable under pullback. Also, since every object of $\mathcal{E}$ is a (κ-filtered) colimit of objects of $\mathcal{C}$, the inclusion $\mathcal{C} \to \mathcal{E}$ is J-dense and hence also K-dense; and every representable functor $\mathcal{E}^{\mathrm{op}} \to \mathbf{Set}$ is a J-sheaf and K-separated, so we have $\mathcal{E} \subseteq \mathbf{Bisep}(\mathcal{C}, J, K)$ as subcategories of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

For the reverse inclusion, we need to show that every (J, K)-biseparated functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ preserves κ-small limits. Consider first a coproduct $U = \sum_{i \in I} U_i$: by (the infinitary analogue of) A1.5.14, this coproduct is quasi-disjoint, i.e. the coprojections $\nu_i \colon U_i \to U$ are monic and $U_i \cap U_j$ is a quasi-initial object for each $i \neq j$. In particular, each $U_i \cap U_j$ is K-covered by the empty sieve; so if $A \colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is (J, K)-biseparated, then each $F(U_i \cap U_j)$ is either empty or a singleton. Now the ν_i generate a J-covering sieve; the sheaf axiom for this sieve says that $A(U)$ is the limit of a diagram whose vertices are the $A(U_i)$ and the $A(U_i \cap U_j)$, but the triviality of the latter means that we can omit them from the diagram without changing its limit. (Note that if any $A(U_i \cap U_j)$ is empty, then both $A(U_i)$ and $A(U_j)$ must be empty too.) So $A(U) \to \prod_{i \in I} A(U_i)$ is an isomorphism, as required.

It remains to consider coequalizers. First we note that the Yoneda embedding $h\colon \mathcal{C} \to \mathbf{Bisep}(\mathcal{C}, J, K)$ preserves regular epimorphisms: since a regular epimorphism $q\colon V \twoheadrightarrow W$ generates a J-covering sieve, $h(q)$ is the coequalizer of its kernel-pair in $\mathbf{Bisep}(\mathcal{C}, J, K)$. Now consider an arbitrary coequalizer diagram

$$U \underset{g}{\overset{f}{\rightrightarrows}} V \xrightarrow{\quad q \quad} W$$

in $\mathcal{C}$; let $q'\colon h(B) \twoheadrightarrow Q$ be the coequalizer of $(h(f), h(g))$ in $\mathbf{Bisep}(\mathcal{C}, J, K)$, and $(a, b)\colon R \rightrightarrows h(V)$ the kernel-pair of q' in this category. We note that $h(q)$ factors through q' (so that $h(q)a = h(q)b$), and that $(h(f), h(g))$ factors through (a, b).

We claim next that R lies in $\mathcal{E}$. We have a monomorphism $(a, b)\colon R \rightarrowtail h(V) \times h(V)$; but $h(V) \times h(V)$ lies in $\mathcal{E}$, so if we are given any other coequalizer diagram $U' \rightrightarrows V' \twoheadrightarrow W'$ in $\mathcal{C}$ and a morphism $h(V') \to R$ having equal composites with $h(U') \rightrightarrows h(V')$, then the composite $h(V') \to R \rightarrowtail h(V) \times h(V)$ factors uniquely through $h(V') \twoheadrightarrow h(W')$, and from the fact that h preserves regular epimorphisms we get a unique lifting of this factorization through $R \rightarrowtail h(V) \times h(V)$. So R preserves equalizers as a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$; but by the argument already given we know that it preserves κ-small products, and hence it lies in $\mathcal{E}$.

In fact $(a, b)\colon R \rightarrowtail h(V) \times h(V)$ is a cocover in $\mathbf{Bisep}(\mathcal{C}, J, K)$, since it is regular monic; but the inclusion $\mathcal{E} \to \mathbf{Bisep}(\mathcal{C}, J, K)$ is easily seen to preserve epimorphisms, so it is also a cocover in $\mathcal{E}$. Hence $R \rightrightarrows h(V)$ is a strong equivalence relation in $\mathcal{E}$, and so it is the kernel-pair of its coequalizer there; but the latter may be identified with $h(q)$, by the existence of the factorizations mentioned earlier and the fact that h preserves coequalizers as a functor $\mathcal{C} \to \mathcal{E}$. So $h(q)$ is isomorphic to q', since both are regular epimorphisms in $\mathbf{Bisep}(\mathcal{C}, J, K)$ and they have the same kernel-pair. Thus we have shown that the inclusion $\mathcal{C} \to \mathbf{Bisep}(\mathcal{C}, J, K)$ preserves coequalizers; equivalently, every object of $\mathbf{Bisep}(\mathcal{C}, J, K)$, considered as a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$, preserves equalizers.

□

A category satisfying the equivalent conditions of 2.2.13 is called a *Grothendieck quasitopos.*

Examples 2.2.14 (*a*) We saw in A2.6.4(*a*) that any Heyting algebra is a quasitopos. Clearly, a Heyting algebra is a Grothendieck quasitopos iff it is complete, i.e. iff it is a frame. Given a frame L, we know from 1.3.15 that L may be recovered (up to equivalence) as the lattice of subterminal objects of $\mathbf{Sh}(L, J)$, where J is the canonical coverage on L; equivalently, it can be represented as $\mathbf{Bisep}(L, J, K)$ where K is the coverage (sometimes called the *chaotic coverage*) in which every sieve is covering, and so the only sheaf is the terminal object.

(*b*) The quasitopos **Fre** of Fréchet spaces (A2.6.4(*c*)) was identified with a full subcategory of the topos $\mathbf{Sh}(M, J)$ of A2.1.11(*j*), where M is the monoid of

continuous endomorphisms of the space $\mathbb{N}^+$ and J is (the Grothendieck coverage generated by) the coverage defined in A2.1.11(j). It is not hard to see that **Fre** consists precisely of the $\neg\neg$-separated objects of this topos; equivalently, we may identify it with $\mathbf{Bisep}(M, J, K)$, where K is the coverage consisting of all jointly-surjective sieves of endomorphisms of $\mathbb{N}^+$ (equivalently, of those sieves which contain all the constant maps $\mathbb{N}^+ \to \mathbb{N}^+$). So it is a Grothendieck quasitopos.

The assumption '$\mathcal{E}$ has a generating set' in 2.2.13(iii) cannot be weakened (as in 2.2.8(iv)) to '$\mathcal{E}$ has a separating set'. The quasitopos **Choq** of Choquet spaces, considered in A2.6.4(b), has a single separator, namely 1; but it does not have a generating set, by an argument like that given for **Sp** in A1.2.5. Similarly, the assumption of quasi-effectiveness cannot be omitted from 2.2.13(iv) (and effectiveness cannot be omitted from 2.2.8(ix)). We shall give a counterexample in 4.2.4 below.

We now revert to the study of toposes. Given a (Grothendieck) topos $\mathcal{E}$, by a *site of definition for* $\mathcal{E}$ we mean a site $(\mathcal{C}, J)$ such that $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, J)$. Of course, there are many possible sites of definition for a given topos; but, at least if we restrict our attention to separated sites, it is possible to characterize them via Proposition 2.2.16 below. Given a small site $(\mathcal{C}, J)$, we shall normally write $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ for the composite of the Yoneda embedding $h\colon \mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ with the associated sheaf functor $L\colon [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Sh}(\mathcal{C}, J)$.

Lemma 2.2.15 *Let $(\mathcal{C}, J)$ be a small site. Then the functor $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ is faithful (resp. full and faithful) iff the coverage J is separated (resp. subcanonical).*

Proof If J is subcanonical, then the Yoneda embedding h factors through the inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, and we may identify l with this factorization; so the fact that it is full and faithful follows from the corresponding property of h. If J is merely separated, then for each $U \in \mathrm{ob}\,\mathcal{C}$ the unit map $h(U) \to l(U)$ is monic, and so composition with it induces an injection

$$\begin{aligned}\mathcal{C}(V, U) \cong [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]\,(h(V), h(U)) &\longrightarrow [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]\,(h(V), l(U)) \\ &\cong \mathbf{Sh}(\mathcal{C})\,(l(V), l(U))\end{aligned}$$

which is clearly the effect of the functor l on morphisms $V \to U$.

Conversely, suppose l is faithful. Then for any two objects U and V of $\mathcal{C}$, we have an injection

$$h(U)(V) = \mathcal{C}\,(V, U) \longrightarrow \mathbf{Sh}(\mathcal{C})\,(l(V), l(U)) \cong [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]\,(h(V), l(U)) \cong l(U)(V);$$

and these injections are natural in V, so they define a monomorphism $h(U) \rightarrowtail l(U)$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. Hence by A4.3.6($a$) $h(U)$ is separated for any U, i.e. J is a separated coverage. If l is full as well as faithful, then the above monomorphism is an isomorphism; so the $h(U)$ are all sheaves, i.e. J is subcanonical. □

It follows from 2.2.15 that, given $\mathcal{E}$, every small separated (resp. subcanonical) site of definition for $\mathcal{E}$ occurs, up to equivalence, as a subcategory (resp. a full subcategory) of $\mathcal{E}$ itself. In the subcanonical case, we can delete the word 'small' from this assertion: although the inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ may not have a left adjoint if $\mathcal{C}$ is not small, it is still true that if J is subcanonical then the Yoneda embedding factors through the inclusion, and so we may identify $\mathcal{C}$ with a full subcategory of $\mathbf{Sh}(\mathcal{C}, J)$. Moreover, it is not hard to see which (full) subcategories of $\mathcal{E}$ give rise to such sites: they are precisely those which satisfy the hypotheses of the Comparison Lemma 2.2.3, when we equip $\mathcal{E}$ with its canonical coverage. Finally, the assertion that a full subcategory $\mathcal{C}$ is dense for the canonical coverage is clearly equivalent to saying that ob $\mathcal{C}$ is a separating family for $\mathcal{E}$. Thus we have established

Proposition 2.2.16 *Let $\mathcal{E}$ be a Grothendieck topos. Then, up to equivalence,*

(i) *the small sites of definition $(\mathcal{C}, J)$ for $\mathcal{E}$ such that J is separated are precisely the sites obtained by equipping small dense subcategories of $\mathcal{E}$ with the coverages induced by the canonical coverage on $\mathcal{E}$; and*

(ii) *the essentially small sites of definition $(\mathcal{C}, J)$ for $\mathcal{E}$ such that J is subcanonical are precisely the sites obtained by equipping full subcategories $\mathcal{C}$ of $\mathcal{E}$ such that* ob $\mathcal{C}$ *is a separating family for $\mathcal{E}$ with the coverages induced by the canonical coverage on $\mathcal{E}$.* □

The 'fundamental theorem of topos theory' (A2.3.2) says that if $\mathcal{E}$ is a topos then so is $\mathcal{E}/A$ for any object A of $\mathcal{E}$. Moreover, it is easy to see that if $\mathcal{E}$ is a Grothendieck topos then so is $\mathcal{E}/A$. The following result is often useful in obtaining a site of definition for $\mathcal{E}/A$.

Lemma 2.2.17 *Let $(\mathcal{C}, J)$ be an essentially small subcanonical site, U an object of $\mathcal{C}$. Define a sieve R on an object $(f\colon V \to U)$ of $\mathcal{C}/U$ to be J_U-covering iff the sieve $\{\Sigma_U(g) \mid g \in R\}$ is J-covering on $V = \Sigma_U(f)$. Then $(\mathcal{C}/U, J_U)$ is an essentially small subcanonical site, and $\mathbf{Sh}(\mathcal{C}/U, J_U) \simeq \mathbf{Sh}(\mathcal{C}, J)/l(U)$.*

Proof It is clear that J_U satisfies conditions (C), (L) and (M) if J does; and if $\mathcal{D}$ is a small dense full subcategory of $\mathcal{C}$, then the full subcategory of $\mathcal{C}/U$ on all objects $(f\colon V \to U)$ with $V \in$ ob $\mathcal{D}$ is (small and) dense. Moreover, J_U is subcanonical: given a J_U-covering sieve R on an object $f\colon V \to U$ of $\mathcal{C}/U$ and another object $g\colon W \to U$, a compatible family of morphisms $(h_k\colon \mathrm{dom}\ k \to g \mid k \in R)$ is in particular a compatible family for the sieve $\{\Sigma_U(k) \mid k \in R\}$, and hence induces a unique morphism $h\colon V \to W$, and the latter is a morphism $f \to g$ in $\mathcal{C}/U$ (that is, $gh = f$) since $\mathcal{C}(-, U)$ is a J-sheaf. The final assertion now follows from 2.2.16(ii), since if we identify $\mathcal{C}$ with a dense full subcategory of $\mathbf{Sh}(\mathcal{C}, J)$ then $\mathcal{C}/U$ will be identified with a dense full subcategory of $\mathbf{Sh}(\mathcal{C}, J)/l(U)$ (and the coverage induced on it will be exactly J_U,

since the forgetful functor $\mathbf{Sh}(\mathcal{C}, J)/l(U) \to \mathbf{Sh}(\mathcal{C}, J)$ preserves and reflects epimorphic families). $\square$

Before leaving this section, we revert briefly to the consideration of those coverages which arise as in 2.2.4(d).

Definition 2.2.18 Let $(\mathcal{C}, J)$ be a site.

(a) We say an object V of $\mathcal{C}$ is (J-)*irreducible* if the only J-covering sieve on V is the maximal sieve M_V.

(b) We say J is *rigid* if, for every $U \in \mathrm{ob}\, \mathcal{C}$, the family of all morphisms from J-irreducible objects to U generates a J-covering sieve.

If J is rigid, then for every $U \in \mathrm{ob}\, \mathcal{C}$ the sieve described in 2.2.18(b) must be the unique smallest J-covering sieve on U; in particular, every object has a smallest J-covering sieve (cf. A4.5.2). (But the latter condition does not imply rigidity: for example, on any dense totally ordered set, such as $\mathbb{Q}$ or $\mathbb{R}$, we have a (Grothendieck) coverage in which each object U has just two covering sieves, namely M_U and the sieve of all morphisms $V \to U$ with $V < U$, and this clearly has no irreducible objects.) Clearly, a coverage which arises as in 2.2.4(d) is rigid, the irreducible objects being (retracts of) objects of the full subcategory $\mathcal{D}$; conversely, if J is a rigid coverage on $\mathcal{C}$, then (provided the full subcategory $\mathcal{D}$ of J-irreducible objects of $\mathcal{C}$ is small) the Comparison Lemma yields an equivalence $\mathbf{Sh}(\mathcal{C}, J) \simeq [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$.

Example 2.2.19 Let $\mathcal{C}$ be a small cartesian category, and $\mathbf{Reg}(\mathcal{C})$ the regularization of $\mathcal{C}$, as defined in A1.3.9. Then the regular coverage J on $\mathbf{Reg}(\mathcal{C})$ is rigid; for we saw in the proof of A1.3.9 that objects of the form $I(A)$, $A \in \mathrm{ob}\, \mathcal{C}$, are projective in $\mathbf{Reg}(\mathcal{C})$ (and hence J-irreducible, since any cover $C \twoheadrightarrow I(A)$ is a split epimorphism); on the other hand, as we noted in A1.3.10(b), every object of $\mathbf{Reg}(\mathcal{C})$ can be covered by one of the form $I(A)$. So in this case the Comparison Lemma tells us that $\mathbf{Sh}(\mathbf{Reg}(\mathcal{C}), J)$ is equivalent to the functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. More generally, if $\mathcal{D}$ is any small regular category with enough projectives (i.e. such that every object can be covered by a projective), then the topos of sheaves for the regular coverage on $\mathcal{D}$ is equivalent to $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, where $\mathcal{C}$ is the full subcategory of projective objects of $\mathcal{D}$.

As an application of the notion of rigidity, we may establish a useful companion for A1.1.10:

Lemma 2.2.20 *For a Grothendieck topos $\mathcal{E}$, the following are equivalent:*

(i) *There exists a small category $\mathcal{C}$ such that $\mathcal{E} \simeq [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.*

(ii) *The canonical coverage on $\mathcal{E}$ is rigid.*

(iii) *$\mathcal{E}$ has a separating set of indecomposable projective objects.*

Proof (i) $\Rightarrow$ (iii) follows easily from (the remarks before) A1.1.10, since the representable functors form a separating set for $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

(iii) $\Rightarrow$ (ii): Suppose A is an indecomposable projective in $\mathcal{E}$. Then, given any epimorphic family $(f_i\colon B_i \to A \mid i \in I)$, at least one f_i must be a split epimorphism; for projectivity allows us to find a splitting $s\colon A \to \coprod_{i\in I} B_i$ for the induced epimorphism, and if we pull back the coprojections along s we obtain a coproduct decomposition of A, which must be trivial. So A is irreducible for the canonical coverage on $\mathcal{E}$; thus (iii) says that every object of $\mathcal{E}$ can be covered by irreducible objects.

(ii) $\Rightarrow$ (i) will be immediate from the Comparison Lemma, provided we can show that the full subcategory of irreducible objects for the canonical coverage is equivalent to a small category. But, if $\mathcal{G}$ is any separating set for $\mathcal{E}$, then for each irreducible object A the family of all morphisms from members of $\mathcal{G}$ to A must contain a split epimorphism; hence A is a retract of some member of $\mathcal{G}$, and there are (up to isomorphism) only a set of such retracts. $\square$

In passing, we remark that Grothendieck toposes with a separating set of projective (but not necessarily indecomposable) objects have been characterized in [531].

Rigid coverages also arise from consideration of finite categories:

Lemma 2.2.21 *If $\mathcal{C}$ is a finite Cauchy-complete category, then every (Grothendieck) coverage on $\mathcal{C}$ is rigid.*

Proof Let J be such a coverage. Since J-covering sieves are closed under finite intersections by the remarks after 2.1.8, any object U of $\mathcal{C}$ has a smallest J-covering sieve, which we shall denote by R_U. We must show that R_U is generated by morphisms whose domains are J-irreducible.

We shall say that a morphism $f\colon V \to U$ is *essential* in a sieve R if ($f \in R$ and) it does not factor nontrivially through any other morphism in R, i.e., whenever we have a factorization $f = gh$ with $g \in R$, then h is a split monomorphism. We claim first that if f is essential in R_U, then V is J-irreducible. For the sieve consisting of all composites gh with $g \in R_U$ and $h \in R_{\mathrm{dom}\ g}$ is J-covering by (L) and so contains (in fact equals) R_U; hence we have a factorization $f = gh$ with $g \in R_U$ and $h \in R_{\mathrm{dom}\ g}$, whence h is split monic (with splitting k, say). Now $h \in k^*(R_V)$ since the latter is a J-covering sieve on dom k = dom g; so $kh = 1_V \in R_V$, i.e. $R_V = M_V$.

Thus we are reduced to showing that R_U is generated by its essential members. In fact this is true for any sieve R on an object U of a finite Cauchy-complete category: to prove it, let $f_0\colon V_0 \to U$ be a member of R. We may choose a sequence of morphisms $f_n\colon V_n \to U$ in R, such that f_m factors through f_n whenever $m \leq n$, by the following rules:

(a) If f_n can be factored as $g_n h_n$ where h_n is split epic but not an isomorphism, choose such a factorization and set $f_{n+1} = g_n$. (Note that any such g_n is necessarily in R.)

(*b*) If (*a*) is impossible but f_n can be factored as $g_n h_n$ where $g_n \in R$ and h_n is not split monic, choose such a factorization and set $f_{n+1} = g_n$.

(*c*) If both (*a*) and (*b*) are impossible, then stop.

Clearly, if the process stops then it does so at a morphism f_n which is essential in R. But since R is finite, it must either stop or repeat itself; i.e., if it does not stop, we can find $n < m$ with $f_n = f_m$. Composing the h_j for $n \leq j < m$ in the appropriate order, we thus obtain for each j in this range a morphism $k_j \colon V_j \to V_j$ with $f_j k_j = f_j$. Now k_j cannot be an isomorphism, for if it were then h_j would be a split monomorphism. However, since V_j has only finitely many endomorphisms some power of k_j must be idempotent; since $\mathcal{C}$ is Cauchy-complete this idempotent splits, and hence f_j factors through some nontrivial split epimorphism with domain V_j, so that h_j must be such an epimorphism. Since this is true for every j, $n \leq j < m$, it follows that k_n is a (split) epimorphism; but this is impossible, since we have just seen that some power of k_n is a non-invertible idempotent. So the process described above must stop; hence f_0 factors through an essential member of R. □

The proof of 2.2.21 easily extends to any category $\mathcal{C}$ such that each slice category $\mathcal{C}/U$, $U \in \text{ob}\,\mathcal{C}$, is equivalent to a finite category (cf. A2.1.5).

Corollary 2.2.22 *Every bounded* $\mathbf{Set}_f$*-topos is (equivalent to one) of the form* $[\mathcal{D}^{\text{op}}, \mathbf{Set}_f]$ *where* $\mathcal{D}$ *is a finite category.*

Proof By the analogue of 2.2.8 with 'small' replaced by 'finite', every such topos $\mathcal{E}$ is equivalent to one of the form $\mathbf{Sh}_f(\mathcal{C}, J)$, the category of sheaves of finite sets on a finite site $(\mathcal{C}, J)$. Moreover, since idempotents split in $\mathbf{Set}_f$, we may assume by A1.1.9 that $\mathcal{C}$ is Cauchy-complete. Then by 2.2.21 J is rigid, so by (the finite analogue of) 2.2.4(*d*) $\mathcal{E}$ is equivalent to $[\mathcal{D}^{\text{op}}, \mathbf{Set}_f]$ where $\mathcal{D}$ is the full subcategory of J-irreducible objects of $\mathcal{C}$. (Note, incidentally, that any retract of a J-irreducible object is J-irreducible, so $\mathcal{D}$ is also Cauchy-complete.) □

There is a stronger 'functorial' version of 2.2.22. By the finite analogue of 2.2.10, any cartesian functor between bounded $\mathbf{Set}_f$-toposes has a left adjoint, and hence any geometric morphism between such toposes is essential. But by the finite version of A4.1.5, an essential geometric morphism between toposes of the form $[\mathcal{C}, \mathbf{Set}_f]$ is induced by a functor between the corresponding finite categories, provided they are Cauchy-complete, and so we have

Corollary 2.2.23 *The assignment* $\mathcal{C} \mapsto [\mathcal{C}, \mathbf{Set}_f]$ *is an equivalence of 2-categories from the 2-category* $\mathfrak{Cau}_f$ *of finite Cauchy-complete categories, functors and natural transformations between them, to the 2-category* $\mathfrak{BTop}/\mathbf{Set}_f$ *of bounded* $\mathbf{Set}_f$*-toposes.* □

Thus, if we replace the base topos **Set** by $\mathbf{Set}_f$, the study of 'Grothendieck toposes' reduces entirely to the study of finite (Cauchy-complete) categories. If we view toposes as generalized spaces, this result may be seen as the natural generalization of the well-known fact that the category of finite T_0-spaces is isomorphic to the category of finite posets.

Suggestions for further reading: Artin et al. [36], Borceux & Pedicchio [156], Bunge [182], Carboni & Mantovani [228], Pedicchio [942].

C2.3 Morphisms of sites

In this section we investigate those functors between (the underlying categories of) sites which give rise to geometric morphisms between their toposes of sheaves. Such functors are traditionally called 'continuous functors', but this name is unsatisfactory, for two reasons. First, it clashes with the common use of 'continuous' in category theory to mean 'limit-preserving' (continuity in this sense is actually more akin to a colimit-preservation property); and secondly, whilst the traditional justification for the name is that continuous functors between sites are the natural generalization of continuous maps between spaces, they actually go in the 'algebraic' rather than the 'geometric' direction, i.e. the opposite direction to the geometric morphisms which they induce. We have therefore chosen to abandon the traditional name, and use the slightly more cumbersome but less misleading term 'cover-preserving'; we shall also use the term 'morphism of sites' for a functor which is cartesian and cover-preserving.

Unless otherwise stated, all coverages in this section will be Grothendieck coverages. For most of the section, we shall also assume that the underlying categories of our sites are cartesian; we shall discuss in 2.3.7 below how the theory needs to be adapted if this hypothesis is dropped.

Definition 2.3.1 Let $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ be (cartesian) sites, and $F\colon \mathcal{C} \to \mathcal{D}$ a functor. We say F is *cover-preserving* if, for any $U \in \operatorname{ob} \mathcal{C}$ and any covering sieve $R \in J(U)$, the family $(F(f) \mid f \in R)$ generates a K-covering sieve on $F(U)$. And we shall say that F is a *morphism of sites* $(\mathcal{C}, J) \to (\mathcal{D}, K)$ if it is cartesian and cover-preserving.

Examples 2.3.2 (*a*) If A and B are frames, made into sites via their canonical coverages, then a morphism of sites $A \to B$ is exactly the same thing as a frame homomorphism, i.e. a function preserving finite meets and arbitrary joins.

(*b*) If $\mathcal{C}$ and $\mathcal{D}$ are regular categories equipped with their regular coverages, then a morphism of sites $\mathcal{C} \to \mathcal{D}$ is the same thing as a regular functor, i.e. one preserving finite limits and covers. Similarly, coherent functors between coherent categories are the same thing as morphisms of sites for their coherent coverages.

(c) Given any small cartesian site $(\mathcal{C}, J)$, let $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ denote the composite of the Yoneda embedding $h\colon \mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ and the associated sheaf functor, as in 2.2.15. Then l is cartesian, since both h and the associated sheaf

functor preserve finite limits; and it preserves covers if $\mathbf{Sh}(\mathcal{C}, J)$ is equipped with its canonical topology, since if $R \in J(U)$, then R (regarded as a subobject of $h(U)$ in $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$) is the union of the images of the morphisms $h(f)$, $f \in R$, and the associated sheaf functor maps the inclusion $R \rightarrowtail h(U)$ to an isomorphism. So l is a morphism of sites. If $\mathcal{C}$ is not small, then we do not in general have an associated sheaf functor $[\mathcal{C}^{\text{op}}, \mathbf{Set}] \to \mathbf{Sh}(\mathcal{C}, J)$; but if $(\mathcal{C}, J)$ is essentially small and J is subcanonical then we shall still write $l \colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ for the factorization of the Yoneda embedding through the inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\text{op}}, \mathbf{Set}]$; and in this case also it is easily seen to be a morphism of sites.

Lemma 2.3.3 *Let $F \colon (\mathcal{C}, J) \to (\mathcal{D}, K)$ be a morphism of sites. Then the induced functor $F^* \colon [\mathcal{D}^{\text{op}}, \mathbf{Set}] \to [\mathcal{C}^{\text{op}}, \mathbf{Set}]$ which sends B to $B \circ F$ restricts to a functor $\mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$.*

Proof We have to show that if B is a K-sheaf on $\mathcal{D}$ then $F^*(B)$ is a J-sheaf on $\mathcal{C}$. So let R be a J-covering sieve on an object U of $\mathcal{C}$, and suppose we have a compatible family of elements $(s_f \in B(F(\text{dom } f)) \mid f \in R)$. Let S be the (K-covering) sieve on FU generated by the images under F of the members of R; it is enough to show that there is a unique compatible family $(t_g \mid g \in S)$ of elements of B such that $t_{F(f)} = s_f$ for all $f \in R$. The uniqueness is clear since the $F(f)$ generate S; for the existence, note that if a morphism $g \colon W \to FU$ of S factors through two different morphisms $F(f) \colon FV \to FU$ and $F(f') \colon FV' \to FU$ in the image of R (say $g = F(f)h = F(f')h'$), then it also factors through $F(f'')$ where

$$\begin{tikzcd} V'' \arrow[r] \arrow[d] \arrow[dr, "f''"] & V \arrow[d, "f"] \\ V' \arrow[r, "f'"] & U \end{tikzcd}$$

is a pullback in $\mathcal{C}$, since F preserves this pullback. So the compatibility of the family $(s_f \mid f \in R)$ implies that we have $B(h)(s_f) = B(h'')(s_{f''}) = B(h')(s_{f'})$; hence we can unambiguously define t_g to be $B(h)(s_f)$ for any factorization $g = F(f)h$ of g through a morphism in the image of R. □

The proof of 2.3.3 should be compared with that of 2.2.2(ii): in the earlier result, we did not assume that the inclusion functor was cartesian (although 2.2.2(i) implied that it was cover-preserving), and to compensate for this we needed the hypothesis that its domain was a dense subcategory of its codomain. On the other hand, even without the assumption of cartesianness, the proof of 2.3.3 shows that composition with a cover-preserving functor maps sheaves to separated functors – a fact which we shall use in the proof of 3.1.21 below.

Note, incidentally, that we did not use the full strength of cartesianness of F in 2.3.3, but only the preservation of pullbacks (equivalently, by A1.2.9, of finite

connected limits). If we were to weaken the definition of 'morphism of sites' in this way, then the argument of 2.3.4 below would show that every such morphism yields a pre-geometric morphism of sheaf toposes, as defined in A4.1.13. For example, in the situation of 2.2.17, the forgetful functor $\Sigma_U : \mathcal{C}/U \to \mathcal{C}$ induces the pre-geometric morphism $\mathbf{Sh}(\mathcal{C}) \to \mathbf{Sh}(\mathcal{C})/l(U)$ whose inverse image is $\Sigma_{l(U)}$ (and this example is 'generic', in that any 'pre-morphism of sites' with codomain $\mathcal{C}$ factors as a morphism of sites followed by a functor of the above form).

Corollary 2.3.4 *In addition to the hypotheses of* 2.3.3, *assume that $\mathcal{C}$ is small and that $(\mathcal{D}, K)$ is essentially small. Then there is a geometric morphism $\mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ whose direct image is F^*.*

Proof If $\mathcal{D}$ is also small, it is easy to see that F^* has a left adjoint $F_!$, namely the composite

$$\mathbf{Sh}(\mathcal{C}) \longrightarrow [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{\varinjlim_F} [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{a} \mathbf{Sh}(\mathcal{D})$$

where a is the associated sheaf functor: and this functor is cartesian, since the left Kan extension $\varinjlim_F$ is cartesian by the argument of A4.1.10, and a is cartesian by the argument sketched after 2.2.6 (or by A4.3.1 and A4.4.6). (In fact, the existence of a left adjoint for F^* could have been deduced immediately from 2.2.10, since F^* preserves all small limits as a functor $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ and hence also as a functor $\mathbf{Sh}(\mathcal{D}) \to \mathbf{Sh}(\mathcal{C})$.)

If $\mathcal{D}$ is not small, we do not in general have the associated sheaf functor available; but we may reduce to the case just considered on replacing $\mathcal{D}$ by a small dense full subcategory $\mathcal{D}'$ which is closed under finite limits and contains the image of F. For if we equip this subcategory with the induced coverage $K_{\mathcal{D}'}$, then F is still a morphism of sites $(\mathcal{C}, J) \to (\mathcal{D}', K_{\mathcal{D}'})$, and so we get a cartesian left adjoint for $F^* : \mathbf{Sh}(\mathcal{D}') \to \mathbf{Sh}(\mathcal{C})$. But $\mathbf{Sh}(\mathcal{D}')$ is equivalent to $\mathbf{Sh}(\mathcal{D})$ by 2.2.3. □

Remark 2.3.5 We note that 2.3.4 can be made functorial: given two morphisms of sites $F, G : (\mathcal{C}, J) \rightrightarrows (\mathcal{D}, K)$, every natural transformation $F \to G$ induces a natural transformation $G^* \to F^*$ (since sheaves are contravariant functors), and hence a geometric transformation $(F^*, F_!) \to (G^*, G_!)$. The converse holds if K is subcanonical; for then $l : \mathcal{D} \to \mathbf{Sh}(\mathcal{D})$ is full and faithful, and it follows easily from the adjunction that we have $F_!(l(U)) \cong l(FU)$ for any object U of $\mathcal{C}$. So any natural transformation $F_! \to G_!$ induces a natural transformation $F \to G$ by restriction along l.

Remark 2.3.6 There is also a converse to 2.3.4: if $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ are small (cartesian) sites, and $F : \mathcal{C} \to \mathcal{D}$ is a cartesian functor such that the geometric morphism $(F^*, \varinjlim_f) : [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ of A4.1.10 restricts to a geometric morphism $\mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$, then F is cover-preserving. For, by A4.3.12, such a restriction exists iff F^* maps K-sheaves to J-sheaves, iff $\varinjlim_f$ sends

j-dense monomorphisms to k-dense monomorphisms (where j and k are the local operators corresponding to J and K as in 2.1.10). Now if R is any J-covering sieve on an object U of $\mathcal{C}$, we may regard it as a j-dense subobject of the representable functor $\mathcal{C}(-, U)$; but $\varinjlim_f$ sends the latter to $\mathcal{D}(-, FU)$, and it is easy to see that $\varinjlim_f(R)$ is (the subfunctor corresponding to) the sieve generated by $\{F(f) \mid f \in R\}$. So this sieve must be K-covering.

Remark 2.3.7 If we wish to work with (small) non-cartesian sites, then the argument of 2.3.4 can still be made to work provided the functor $\varinjlim_F : [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is cartesian – for which it suffices that, for each object V of $\mathcal{D}$, the comma category $(V \downarrow F)$ is cofiltered (cf. A4.1.10 and B3.2.8(c)). A functor with this property is commonly called *flat*; in the case when $\mathcal{C}$ has finite limits, the condition is equivalent to saying that F preserves them, since we can recover F as the restriction of $\varinjlim_F$ to the subcategories of representable functors. Moreover, if this condition is satisfied, then the argument of 2.3.3 can also be made to work: in the notation of the proof of 2.3.3, we simply take (V'', h'') to be an object of $(W \downarrow F)$ admitting morphisms in this category to (V, h) and (V', h'), whose composites with f and f' are equal. Accordingly, we define a *morphism of sites* $F: (\mathcal{C}, J) \to (\mathcal{D}, K)$ between not-necessarily-cartesian sites to be a functor which preserves covers and is flat; and the results of this section can all be made to work in this more general context. Nevertheless, for the time being we shall continue to restrict our attention to cartesian sites in what follows.

Given two sites $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$, we cannot in general expect every geometric morphism $f: \mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ to be induced as in 2.3.4 by a morphism of sites $(\mathcal{C}, J) \to (\mathcal{D}, K)$: clearly, a necessary condition for this is that the inverse image functor f^* should map objects in the image of l to objects in the image of l. However, this is essentially the only obstruction:

Lemma 2.3.8 *Let $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ be cartesian sites such that $\mathcal{C}$ is small and $(\mathcal{D}, K)$ is essentially small and subcanonical. Then a geometric morphism $f: \mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ is induced by a morphism of sites $(\mathcal{C}, J) \to (\mathcal{D}, K)$ iff the composite $f^* l: \mathcal{C} \to \mathbf{Sh}(\mathcal{D}, K)$ factors through the embedding $l: \mathcal{D} \to \mathbf{Sh}(\mathcal{D}, K)$.*

Proof It suffices to show that the resulting factorization $F: \mathcal{C} \to \mathcal{D}$ is a morphism of sites: it will then follow immediately that f^* is isomorphic to $F_!$, since the two functors agree on the image of l and both preserve colimits. Clearly, F is cartesian since $f^* l$ is. Also, $f^* l$ preserves covers (that is, maps them to epimorphic families in $\mathbf{Sh}(\mathcal{D}, K)$); and since K is subcanonical it follows from 2.2.16 that the coverage K is that induced on $\mathcal{D}$ by the canonical coverage on $\mathbf{Sh}(\mathcal{D}, K)$, i.e. that $l: \mathcal{D} \to \mathbf{Sh}(\mathcal{D}, K)$ reflects covers as well as preserving them. Hence F preserves covers. □

Corollary 2.3.9 *For any small cartesian site $(\mathcal{C}, J)$ and any Grothendieck topos $\mathcal{E}$, the category $\mathfrak{Top}(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, J))$ of geometric morphisms $\mathcal{E} \to \mathbf{Sh}(\mathcal{C}, J)$ is*

equivalent to the category $\mathfrak{Site}((\mathcal{C}, J), (\mathcal{E}, C))$ *of morphisms of sites and natural transformations between them, where* C *is the canonical coverage on* $\mathcal{E}$.

Proof Since $l: \mathcal{E} \to \mathbf{Sh}(\mathcal{E}, C)$ is an equivalence by 2.2.7, every geometric morphism $\mathcal{E} \to \mathbf{Sh}(\mathcal{C}, J)$ satisfies the hypothesis of 2.3.8. The fact that the assignment $F \mapsto (F^*, F_!)$ yields an equivalence of categories then follows from 2.3.5. □

If we wish to consider only small sites, then we have an alternative to 2.3.9:

Corollary 2.3.10 *For any geometric morphism* $f: \mathcal{F} \to \mathcal{E}$ *between Grothendieck toposes, there exist (small) standard sites of definition* $(\mathcal{C}, J)$ *and* $(\mathcal{D}, K)$ *for* $\mathcal{E}$ *and* $\mathcal{F}$ *respectively, such that* f *is induced by a morphism of sites* $(\mathcal{C}, J) \to (\mathcal{D}, K)$.

Proof First choose a standard site $(\mathcal{C}, J)$ for $\mathcal{E}$, as in 2.2.16; then choose a similar site for $\mathcal{F}$ ensuring that ob $\mathcal{D}$ contains, in addition to a separating set for $\mathcal{F}$, all objects of the form $f^*(l(U))$, $U \in$ ob $\mathcal{C}$. The result is then immediate from 2.3.8. □

Remark 2.3.11 In general, the site $(\mathcal{D}, K)$ in 2.3.10 will depend on the morphism f; we cannot expect to find a single (small) site which will suffice for all geometric morphisms $\mathcal{F} \to \mathcal{E}$. However, there is an important special case in which this is possible: namely, when the topos $\mathcal{E}$ is localic (over **Set**). For in this case we can take $\mathrm{Sub}_{\mathcal{E}}(1)$, with its canonical coverage, as our site for $\mathcal{E}$, and since any inverse image functor preserves 1 and monomorphisms we can take $\mathcal{D}$ to be any small subcategory of $\mathcal{F}$ containing a separating set and (representatives of) all subobjects of 1 in $\mathcal{F}$. In the particular case when $\mathcal{F}$ is also localic, we recover the result of 1.4.5 that $\mathfrak{Top}(\mathcal{F}, \mathcal{E})$ is equivalent to the poset of frame homomorphisms $\mathrm{Sub}_{\mathcal{E}}(1) \to \mathrm{Sub}_{\mathcal{F}}(1)$, by 2.3.2(*a*).

Given a functor $F: \mathcal{C} \to \mathcal{D}$ between two categories and a coverage on one of them, it is natural to ask whether there is a 'best possible' coverage on the other which makes the functor cover-preserving. In one direction this is easy; in the other rather less so.

Lemma 2.3.12 *Given a functor* $F: \mathcal{C} \to \mathcal{D}$ *and a coverage* J *on* $\mathcal{C}$, *there is a unique smallest coverage* $F_\bullet J$ *on* $\mathcal{D}$ *for which* F *is cover-preserving. If* $\mathcal{C}$ *and* $\mathcal{D}$ *are small and cartesian and* F *preserves finite limits, then this coverage makes the square*

$$\begin{tikzcd} \mathbf{Sh}(\mathcal{D}, F_\bullet J) \arrow[r] \arrow[d] & \mathbf{Sh}(\mathcal{C}, J) \arrow[d] \\ {[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]} \arrow[r, "f"] & {[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]} \end{tikzcd}$$

a pullback in $\mathfrak{Top}$, *where* f *is the morphism induced by* F *as in* A4.1.10.

Proof Since an intersection of (Grothendieck) coverages is a coverage, we simply define $F_{\bullet}J$ to be the intersection of all coverages containing the sieve generated by $\{F(f) \mid f \in R\}$ for every J-covering sieve R. For the second part, we note that $\mathbf{Sh}(\mathcal{D}, F_{\bullet}J)$ is the largest subtopos of $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ for which the composite $\mathbf{Sh}(\mathcal{D}) \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ factors through $\mathbf{Sh}(\mathcal{C}, J)$, by 2.3.6; but this was how we defined pullbacks of inclusions in $\mathfrak{Top}$, in A4.5.14(e). □

Lemma 2.3.13 *Suppose given a functor $F\colon \mathcal{C} \to \mathcal{D}$ and a coverage K on $\mathcal{D}$. If $\mathcal{C}$ has and F preserves pullbacks, then there is a unique largest coverage $F^{\bullet}K$ on $\mathcal{C}$ for which F is cover-preserving. If $\mathcal{C}$ and $\mathcal{D}$ are small and cartesian and F preserves finite limits, then $\mathbf{Sh}(\mathcal{C}, F^{\bullet}K) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is the image (in the sense of* A4.2.10*) of the composite $\mathbf{Sh}(\mathcal{D}, K) \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.*

Proof We define $F^{\bullet}K(U)$ to consist of all sieves R on U for which the set $\{Ff \mid f \in R\}$ generates a K-covering sieve on FU. Clearly, if this is indeed a coverage then it is the largest coverage for which F is cover-preserving. It is straightforward to verify that $F^{\bullet}K$ satisfies (M) and (L) if K does; the difficulty lies in verifying (C) (or equivalently (C')), since if $g\colon U' \to U$ is a morphism of $\mathcal{C}$ then the sieve generated by $\{Ff' \mid f' \in g^*R\}$ is in general smaller than the pullback along Fg of the sieve generated by $\{Ff \mid f \in R\}$. However, under the hypothesis that F preserves pullbacks, the two are equal, since both are generated by the left vertical morphisms in the pullback squares

$$\begin{tikzcd} FV' \arrow[r] \arrow[d] & FV \arrow[d, "Ff"] \\ FU' \arrow[r, "Fg"] & FU \end{tikzcd}$$

as f ranges over the morphisms in R. The second assertion again follows from 2.3.6, since $\mathbf{Sh}(\mathcal{C}, F^{\bullet}K)$ is the smallest subtopos of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ through which the given composite factors. □

Combining 2.3.9 and 2.3.12, we may obtain a description of weighted limits in $\mathfrak{BTop}/\mathbf{Set}$ (cf. Section B4.1) in terms of sites.

Lemma 2.3.14 *Suppose given a small weighted diagram (D, W) in $\mathfrak{BTop}/\mathbf{Set}$; suppose further that we have chosen cartesian sites $(\mathcal{C}_i, J_i)$ for each topos $\mathcal{E}_i$ appearing as a vertex of this diagram, in such a way that each geometric morphism $f_\alpha\colon \mathcal{E}_j \to \mathcal{E}_i$ appearing as an edge of the diagram is induced by a morphism of sites $F_\alpha\colon (\mathcal{C}_i, J_i) \to (\mathcal{C}_j, J_j)$. Then the weighted limit of the diagram may be identified with $\mathbf{Sh}(\mathcal{C}_\infty, J_\infty)$, where $\mathcal{C}_\infty$ is the weighted colimit of the diagram formed by the $\mathcal{C}_i$ and the F_α in the 2-category $\mathfrak{Cart}$ of small cartesian categories and cartesian functors, and J_∞ is the smallest coverage on $\mathcal{C}_\infty$ for which each leg $F_i\colon \mathcal{C}_i \to \mathcal{C}_\infty$ of the colimit cone is cover-preserving (equivalently, J_∞ is the join of the coverages $(F_i)_{\bullet}J_i$ in the lattice of (Grothendieck) coverages on $\mathcal{C}_\infty$).*

Proof By 2.3.9, a weighted cone over D with vertex $\mathcal{F}$ corresponds to a weighted cone in $\mathfrak{Cart}$ under the diagram formed by the $\mathcal{C}_i$ and F_α, all of whose legs are cover-preserving functors. If we impose on $\mathcal{C}_\infty$ the coverage J_∞ described in the statement, then a cartesian functor $G: \mathcal{C}_\infty \to \mathcal{F}$ is cover-preserving (for the canonical coverage C on $\mathcal{F}$) iff $J_\infty \subseteq G^\bullet(C)$, iff each composite GF_i is cover-preserving. So the result is immediate. □

Lemma 2.3.14 could be formulated for sites whose underlying categories are not cartesian, provided we consider weighted colimits in an appropriate 2-category $\mathfrak{Flat}$ of small categories and flat functors between them. Of course, the usefulness of the result depends on our ability to give explicit descriptions of the required weighted colimits, and this tends to be simpler in $\mathfrak{Cart}$ than in $\mathfrak{Flat}$. There are a number of particular cases where we can do so relatively easily.

Examples 2.3.15 (*a*) Finite coproducts in $\mathfrak{Cart}$ coincide with products, in much the same way that they coincide in an abelian category: given cartesian categories $\mathcal{C}_1$ and $\mathcal{C}_2$, the projections $P_i: \mathcal{C}_1 \times \mathcal{C}_2 \to \mathcal{C}_i$ have (automatically cartesian!) right adjoints R_i (defined by $R_1(U) = (U, 1)$ and $R_2(V) = (1, V)$), which have the universal property of a coproduct cone. (Given cartesian functors $G_1: \mathcal{C}_1 \to \mathcal{D}$ and $G_2: \mathcal{C}_2 \to \mathcal{D}$, the induced cartesian functor $\mathcal{C}_1 \times \mathcal{C}_2 \to \mathcal{D}$ sends an object (U, V) to $G_1U \times G_2V$.) So 2.3.14 yields a description of the product of two Grothendieck toposes $\mathbf{Sh}(\mathcal{C}_1, J_1)$ and $\mathbf{Sh}(\mathcal{C}_2, J_2)$ in the form $\mathbf{Sh}(\mathcal{C}_1 \times \mathcal{C}_2, J_3)$ for a suitable coverage J_3. In particular, restricting to the case when J_1 and J_2 are trivial, we recover (part of) the result of B3.2.13 that $[\mathcal{C}_1^{\mathrm{op}}, \mathbf{Set}] \times [\mathcal{C}_2^{\mathrm{op}}, \mathbf{Set}]$ is equivalent to $[(\mathcal{C}_1 \times \mathcal{C}_2)^{\mathrm{op}}, \mathbf{Set}]$. (In B3.2.13 we proved this without the restriction that $\mathcal{C}_1$ and $\mathcal{C}_2$ should be cartesian; but we cannot do the same here, because the coincidence of finite products and coproducts fails to hold in $\mathfrak{Flat}$.)

(*b*) There are other instances of limit–colimit coincidence in $\mathfrak{Cart}$, which can be exploited in the same way. For example, suppose given a cartesian functor $F: \mathcal{C}_1 \to \mathcal{C}_2$; then the comma category $\mathbf{Gl}(F) = (\mathcal{C}_2 \downarrow F)$ (cf. A2.1.12) has the property that its two projections $P_i: \mathbf{Gl}(F) \to \mathcal{C}_i$ have right adjoints R_i (specifically, $R_1(U) = (U, FU, 1_{FU})$ and $R_2(V) = (1, V, (V \to 1))$), and there is an obvious natural transformation $\beta: R_1 \to R_2F$ (the mate of the universal 2-cell $\alpha: P_2 \to FP_1$) giving $\mathbf{Gl}(F)$ the universal property of a cocomma object $(\mathcal{C}_1 \uparrow F)$ in $\mathfrak{Cart}$. Given cartesian functors $G_i: \mathcal{C}_i \to \mathcal{D}$ $(i = 1, 2)$ and a natural transformation $\gamma: G_1 \to G_2F$, the induced cartesian functor $\mathbf{Gl}(F) \to \mathcal{D}$ sends (U, V, f) to the pullback of

$$\begin{tikzcd} & G_1U \arrow[d, "\gamma_U"] \\ G_2V \arrow[r, "G_2f"] & G_2FU \end{tikzcd} .$$

(Observant readers will recognize that we have made use of this property of $\mathbf{Gl}(F)$ several times in earlier parts of this book, when applying the glueing construction to toposes.) Thus, if we wish to construct a site of definition for a comma object of the form $(\mathcal{E}_1 \downarrow f)$ in $\mathfrak{BTop}/\mathbf{Set}$, for a geometric morphism $f: \mathcal{E}_2 \to \mathcal{E}_1$, we need only find cartesian sites of definition $(\mathcal{C}_i, J_i)$ for the $\mathcal{E}_i$ such that f is induced by a morphism of sites $F: (\mathcal{C}_1, J_1) \to (\mathcal{C}_2, J_2)$, and then impose the appropriate coverage on $\mathbf{Gl}(F)$.

(*c*) Coinverters in $\mathfrak{Cart}$ may be constructed using categories of fractions in the sense of Gabriel and Zisman [387]. To form the coinverter of a natural transformation $\alpha: F \to G$ between cartesian functors $\mathcal{B} \rightrightarrows \mathcal{C}$, let Σ be the set of those morphisms of $\mathcal{C}$ which can be expressed as composites of pullbacks of components $\alpha_U: FU \to GU$ of α. Then Σ admits a calculus of right fractions, so that the category of fractions $\mathcal{C}[\Sigma^{-1}]$ is cartesian, and the canonical functor $P: \mathcal{C} \to \mathcal{C}[\Sigma^{-1}]$ preserves finite limits; on the other hand, if $H: \mathcal{C} \to \mathcal{D}$ is any cartesian functor such that $H \circ \alpha$ is an isomorphism, then H must send all morphisms in Σ to isomorphisms, and so it factors uniquely through P. Thus we obtain a description of inverters in $\mathfrak{BTop}/\mathbf{Set}$ in terms of sites (though we could alternatively have done this by retaining the same category $\mathcal{C}$ but imposing a larger coverage on it; cf. B4.1.7).

(*d*) By putting together the last three constructions, we may obtain a description of general cocomma objects in $\mathfrak{Cart}$, as follows. Suppose given a span

$$\mathcal{C}_2 \xleftarrow{\ F\ } \mathcal{C}_1 \xrightarrow{\ G\ } \mathcal{C}_3$$

in $\mathfrak{Cart}$. We first form the comma category $\mathbf{Gl}(F,G) = (\mathcal{C}_2 \times \mathcal{C}_3 \downarrow (F,G))$, whose objects are quintuples (U, V, W, f, g) such that $f: V \to FU$ and $g: W \to GU$. As in (*b*), this has the universal property of a cocomma object $(\mathcal{C}_1 \uparrow (F,G))$ in $\mathfrak{Cart}$, but since $\mathcal{C}_2 \times \mathcal{C}_3$ is also a coproduct in $\mathfrak{Cart}$ we may regard this as a 'lax pushout', that is a normalized lax colimit for the given span. Explicitly, this means that we have functors $R_i: \mathcal{C}_i \to \mathbf{Gl}(F,G)$ $(i = 1,2,3)$ together with 2-cells $\alpha: R_1 \to R_2F$ and $\beta: R_1 \to R_3G$, which are universal in the obvious sense. (The word 'normalized' signifies the fact that, in comparison with the lax and oplax limits considered in B1.1.6, we have not required the presence in the lax cone of 2-cells corresponding to the identity morphisms in the diagram – or, if you prefer, that we have required these 2-cells to be identities.)

Now, if we wish to construct $(F \uparrow G)$ in $\mathfrak{Cart}$, we have simply to form the coinverter $L: \mathbf{Gl}(F,G) \to \mathcal{L}$ of the 2-cell α, since then the 2-cell $(L \circ \beta)(L \circ \alpha)^{-1}$ will have the appropriate universal property. As we observed in (*c*), we may do this by forming a category of fractions; in the present case, we have to form $\mathbf{Gl}(F,G)[\Sigma^{-1}]$, where Σ consists of all morphisms

$(u, v, w)\colon (U, V, W, f, g) \to (U', V', W', f', g')$ such that v is an isomorphism and

$$\begin{tikzcd} W \arrow[r, "g"] \arrow[d, "w"] & GU \arrow[d, "Gu"] \\ W' \arrow[r, "g'"] & GU' \end{tikzcd}$$

is a pullback. For every such morphism may be expressed as a pullback of the U-component $((U \to 1), 1, (GU \to 1))\colon (U, FU, GU, 1, 1) \to (1, FU, 1, (FU \to 1), 1)$ of α; and the class of such morphisms is stable under composition.

Using the explicit description of cocomma objects in $\mathfrak{Cart}$ provided by 2.3.15(d), we may prove a result promised in Section B4.1. First we need

Lemma 2.3.16 *Suppose given a cocomma square*

$$\begin{tikzcd} \mathcal{C}_1 \arrow[r, "F"] \arrow[d, "G"] \arrow[dr, phantom, "\downarrow"] & \mathcal{C}_2 \arrow[d, "P"] \\ \mathcal{C}_3 \arrow[r, "Q"] & \mathcal{D} \end{tikzcd}$$

in $\mathfrak{Cart}$, where F has a (not necessarily cartesian) left adjoint L. Then Q has a $\mathcal{C}_3$-indexed left adjoint M, and the Beck–Chevalley natural transformation $MP \to GL$ is an isomorphism.

Proof By 2.3.15(d), we may identify $\mathcal{D}$ with a category of fractions $\mathbf{Gl}(F, G)[\Sigma^{-1}]$; but in the present case we may also identify it with a subcategory of $\mathbf{Gl}(F, G)$, namely the full subcategory on those objects (U, V, W, f, g) for which the transpose $\overline{f}\colon LV \to U$ of f is an isomorphism. For this subcategory is coreflective in $\mathbf{Gl}(F, G)$; the coreflector sends an arbitrary object (U, V, W, f, g) to $(LV, V, \widehat{W}, \eta_V, \hat{g})$, where η is the unit of $(L \dashv F)$ and

$$\begin{tikzcd} \widehat{W} \arrow[r, "\hat{f}"] \arrow[d, "\hat{g}"] & W \arrow[d, "g"] \\ GLV \arrow[r, "G\overline{f}"] & GU \end{tikzcd}$$

is a pullback. The counit $(\overline{f}, 1, \hat{f})\colon (LV, V, \widehat{W}, \eta_V, \hat{g}) \to (U, V, W, f, g)$ is easily seen to belong to Σ, and the coreflector maps all morphisms in Σ to isomorphisms, from which it follows easily that it has the universal property of the canonical functor $\mathbf{Gl}(F, G) \to \mathbf{Gl}(F, G)[\Sigma^{-1}]$.

We may further simplify our description of $\mathcal{D}$ by considering its objects to be triples (V, W, g) where $g\colon W \to GLV$ in $\mathcal{C}_3$; in other words, it is simply the comma category $(\mathcal{C}_3 \downarrow GL)$. In terms of this description, the functors P and Q are given by

$$P(V) = (V, GLV, 1_{GLV}) \quad \text{and} \quad Q(W) = (1, GL(1) \times W, \pi_1)\,.$$

It is now easy to see that both P and Q have left adjoints given by the two projections, and that the left adjoint M of Q satisfies $MP(V) \cong GL(V)$ for any $V \in \mathrm{ob}\mathcal{C}_2$. Moreover, the product of two objects (V, W, g) and (V', W', g') of $\mathcal{D}$ is of the form $(V \times V', W'', g'')$, where W'' is the limit of the diagram

$$\begin{tikzcd}
W \arrow[dr, "g"'] & & GL(V \times V') \arrow[dl, "GL\pi_1"'] \arrow[dr, "GL\pi_2"] & & W' \arrow[dl, "g'"] \\
& GLV & & GLV' &
\end{tikzcd},$$

from which it follows easily that the Frobenius reciprocity condition

$$M((V, W, g) \times Q(W')) \cong M(V, W, g) \times W'$$

holds. Using this (and the analogous result for pullbacks in $\mathcal{D}$), we may easily verify that the left adjoints M^W for the functors $Q/W\colon \mathcal{C}_3/W \to \mathcal{D}/QW$, defined as in B1.2.4, fit together to form a $\mathcal{C}_3$-indexed left adjoint for Q. □

Theorem 2.3.17 *Suppose given a comma square*

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "q"] \arrow[d, "p"'] \arrow[dr, phantom, "\rightarrow"] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"'] & \mathcal{E}
\end{tikzcd}$$

in $\mathfrak{BTop}/\mathbf{Set}$, *where f is essential. Then q is $\mathcal{G}$-essential (that is, q^* has a $\mathcal{G}$-indexed right adjoint), and the Beck–Chevalley natural transformation $q_! p^* \to g^* f_!$ is an isomorphism.*

Proof We need to choose small full subcategories $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ of $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$ respectively, such that each is closed under finite limits and contains a generating set for the relevant topos, and such that g^*, f^* and $f_!$ all restrict to functors between the corresponding subcategories. To do this, we take $\mathcal{A}$ and $\mathcal{B}$ to be the unions of increasing sequences $\mathcal{A}_n$ and $\mathcal{B}_n$ respectively, chosen such that f^* maps $\mathcal{A}_n$ into $\mathcal{B}_n$ and $f_!$ maps $\mathcal{B}_n$ into $\mathcal{A}_{n+1}$ for each n; then we choose $\mathcal{C}$ to be any small full

subcategory of $\mathcal{G}$ which contains the image of $\mathcal{A}$ under g^*, as well as a generating set, and is closed under finite limits.

Now let F, G and L denote the restrictions of f^*, g^* and $f_!$ to the relevant small subcategories, and form the cocomma object $\mathcal{D} = (F \uparrow G)$ in $\mathfrak{Cart}$ as in 2.3.16 – that is, we take $\mathcal{D}$ to be the comma category $(\mathcal{C} \downarrow GL)$ in the classical sense. By 2.3.14, the comma object $(f \downarrow g)$ in $\mathfrak{BTop}/\mathbf{Set}$ may be taken to be $\mathbf{Sh}(\mathcal{D}, K)$, where K is the smallest coverage on $\mathcal{D}$ for which the functors P and Q are cover-preserving (relative to the coverages H and J on $\mathcal{B}$ and $\mathcal{C}$ induced by the canonical coverages on $\mathcal{F}$ and $\mathcal{G}$).

We claim that K may be described explicitly as follows: a family

$$((v_i, w_i) \colon (V_i, W_i, g_i) \to (V, W, g) \mid i \in I)$$

generates a K-covering sieve iff (i) $(w_i \mid i \in I)$ generates a J-covering sieve, and (ii) if I is inhabited, then $(v_i \mid i \in I)$ generates an H-covering sieve. For the generating covers all satisfy this condition (note that GL is cover-preserving, since it is the restriction of a functor $\mathcal{F} \to \mathcal{G}$ which preserves epimorphic families); the covers satisfying the condition form a Grothendieck coverage; and every cover satisfying the condition may be obtained by pulling back and composing covers which are the images of H-covering or J-covering families. In particular, it follows that the projection $M \colon \mathcal{D} \to \mathcal{C}$ (the left adjoint of Q) is cover-preserving.

Also, although K is not subcanonical (the object $(1, 0, 0 \to GL(1))$ is covered by the empty sieve, though it is not strict initial unless $\mathcal{F}$ is degenerate), it is easy to see that any representable functor of the form $\mathcal{D}(-, QW)$ is a K-sheaf. Hence we may simplify the description of q^*, at least on objects of $\mathcal{G}$ which belong to the subcategory $\mathcal{C}$ (equivalently, are representable as sheaves on $\mathcal{C}$): by definition, $q^*(\mathcal{C}(-, W))$ is obtained by computing the left Kan extension of $\mathcal{C}(-, W)$ along Q and then applying the associated sheaf functor, but the Kan extension is simply $\mathcal{D}(-, QW) \cong \mathcal{C}(M-, W)$, so it is a sheaf already. It follows that if, for an arbitrary sheaf D on $(\mathcal{D}, K)$, we define $q_!(D)$ to be the associated J-sheaf of the left Kan extension $\varinjlim_M(D)$, then morphisms $q_! D \to C$ correspond naturally to morphisms $D \to q^* C$ for all representable sheaves C on $(\mathcal{C}, J)$. But since every object of $\mathcal{G} \simeq \mathbf{Sh}(\mathcal{C}, J)$ can be J-covered by representables (or, if you prefer, because we could have included any given object of $\mathcal{G}$ in our site – indeed, we could really have ignored questions of size and taken $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ to be the whole of $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$ respectively), it follows that $q_!$ is the desired left adjoint to q^*. Further, the Beck–Chevalley condition $q_! p^* \cong g^* f_!$, and the fact that the adjunction $(q_! \dashv q^*)$ is $\mathcal{G}$-indexed, follow from the corresponding parts of 2.3.16 and the fact that any given object of $\mathcal{F}$ (resp. any given morphism of $\mathcal{G}$) could have been included in our generating subcategory $\mathcal{B}$ (resp. $\mathcal{C}$). □

The arguments in the proof of 2.3.17 are all constructive, and so can be applied to comma squares in $\mathfrak{BTop}/\mathcal{S}$ for an arbitrary topos $\mathcal{S}$ with a natural number object – provided we strengthen the hypothesis 'f is essential' to 'f is $\mathcal{S}$-essential' (cf. B3.1.1). Thus we may obtain Pitts's theorem as stated in B4.1.8.

Given sites $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$, we say a functor $L: \mathcal{D} \to \mathcal{C}$ is *cover-reflecting* if, given any object V of $\mathcal{D}$ and any J-covering sieve R on LV, there exists a K-covering sieve S on V such that $L(f) \in R$ for all $f \in S$ – equivalently, the sieve of all $f: U \to V$ such that $L(f) \in R$ is in K. (Authors who use the term 'continuous' for cover-preserving functors normally call functors with this property 'cocontinuous'; the term 'comorphism of sites' is also in use.)

Proposition 2.3.18 *Let $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ be small sites, and $L: \mathcal{D} \to \mathcal{C}$ a functor. The following conditions are equivalent:*

(i) *L is cover-reflecting.*

(ii) *There is a geometric morphism $g: \mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ whose inverse image is the composite*

$$\mathbf{Sh}(\mathcal{C}, J) \xrightarrow{i} [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{L^*} [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \xrightarrow{a} \mathbf{Sh}(\mathcal{D}, K)$$

(where i is the inclusion functor and a is the associated sheaf functor).

(iii) *The right Kan extension functor $\varprojlim_L: [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ maps K-sheaves to J-sheaves.*

(iv) *$L^*: [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ maps j-dense monomorphisms to k-dense monomorphisms, where j and k are the local operators corresponding to J and K as in* 2.1.10.

Proof The composite in (ii) is cartesian, since each of the three factors is; so it is an inverse image functor iff it has a right adjoint, which can only be the restriction of $\varprojlim_L$ to sheaves. Thus (ii) and (iii) are equivalent; and the equivalence of (iii) and (iv) was shown in A4.3.12. So it suffices to show that (i) and (iv) are equivalent.

First assume (i); let $A' \rightarrowtail A$ be a j-dense monomorphism in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. We have to show that, for each $x \in L^*A(V) = A(LV)$, the sieve of morphisms $f: V' \to V$ such that $L^*A(f)(x) \in L^*A'(V')$ is K-covering. But this is exactly $\{f \mid L(f) \in R\}$, where R is the J-covering sieve $\{g: U \to LV \mid A(g)(x) \in A'(U)\}$.

Conversely, assume (iv), and let R be a J-covering sieve on an object of the form LV. We may regard R as a j-dense subobject of $\mathcal{C}(-, LV)$, and form the pullback

$$\begin{tikzcd} S \arrow[r] \arrow[d, tail] & L^*R \arrow[d, tail] \\ \mathcal{D}(-, V) \arrow[r, "\eta"] & L^*(\mathcal{C}(-, LV)) \end{tikzcd}$$

where η is the unit of $(\varinjlim_L \dashv L^*)$, in other words $\eta_{V'}(f: V' \to V) = L(f)$. Then $S \rightarrowtail \mathcal{D}(-, V)$ is k-dense; but the sieve corresponding to S is precisely $\{f \mid L(f) \in R\}$. So L is cover-reflecting. $\square$

There is a 'cover-reflecting' version of 2.3.10: for any geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between Grothendieck toposes, we can find sites of definition for $\mathcal{E}$ and $\mathcal{F}$ such that f is induced by a cover-reflecting functor between them. However, the proof of this fact requires consideration of internal sites, and so we defer it to 2.5.5 below. On the other hand, we may immediately establish the cover-reflecting analogues of 2.3.12 and 2.3.13:

Lemma 2.3.19 *Let $L: \mathcal{D} \to \mathcal{C}$ be a functor between small categories.*

(i) *For each coverage J on $\mathcal{C}$, there is a unique smallest coverage L^+J on $\mathcal{D}$ for which L is cover-reflecting.*

(ii) *For each coverage K on $\mathcal{D}$, there is a unique largest coverage L_+K on $\mathcal{C}$ for which L is cover-reflecting.*

Proof (i) One might be tempted to define $L^+J(V)$ to consist of all sieves of the form $\{f: V' \to V \mid Lf \in R\}$ as R ranges over $J(LV)$. It is easy to verify that this definition satisfies (C') and (M) if J does; but in general it fails to satisfy (L), so we actually have to take L^+J to be the Grothendieck coverage generated by this collection of sieves, as in 2.1.9. Alternatively, we could define L^+J to be the unique (Grothendieck) coverage on $\mathcal{D}$ making the square

$$\begin{tikzcd}
\mathbf{Sh}(\mathcal{D}, L^+J) \arrow[r] \arrow[d] & \mathbf{Sh}(\mathcal{C}, J) \arrow[d] \\
{[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]} \arrow[r] & {[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]}
\end{tikzcd}$$

a pullback (where the bottom edge is the geometric morphism induced by L as in A4.1.4), and then appeal to 2.3.18.

(ii) Here we can give an explicit description of L_+K (though, once again, we could also define it as the coverage corresponding to the image of the composite $\mathbf{Sh}(\mathcal{D}, K) \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$): a sieve R on an object U of $\mathcal{C}$ is L_+K-covering iff, for all $f: LV \to U$ in $\mathcal{C}$, the sieve $\{g: V' \to V \mid Lg \in f^*R\}$ is in $K(V)$. Clearly, if J is any coverage making L cover-reflecting, then any J-covering sieve must satisfy this condition; so L_+K will be the largest such coverage provided it is indeed a coverage. The verification of conditions (C') and (M) is easy; let us verify (L). Suppose given an L_+K-covering sieve R on U, and another sieve S on U such that f^*S is L_+K-covering for each $f \in R$. For any $h: LV \to U$, let R_h denote the sieve $\{g \mid Lg \in h^*R\}$ on V, and similarly let $S_h = \{g \mid Lg \in h^*S\}$. Then R_h is K-covering, and for each $g: V' \to V$ in R_h the composite $h \circ Lg$ is in R, so $(LG)^*h^*S$ is in $L_+K(LV')$. But this implies that $\{g': V'' \to V' \mid Lg' \in (Lg)^*h^*S\}$ is in $K(V')$; and this sieve is exactly $g^*(S_h)$. So by condition (L) for K we deduce $S_h \in K(V)$; since this is true for all h, we deduce $S \in L_+K(U)$. $\square$

As an application of 2.3.19, we give a description of the localic reflection (cf. A4.6.12) of a Grothendieck topos in terms of sites:

Example 2.3.20 Let $(\mathcal{C}, J)$ be an arbitrary small site, and let $Q: \mathcal{C} \to \mathcal{P}$ be the preorder reflection of $\mathcal{C}$, i.e. the quotient of $\mathcal{C}$ by the congruence which identifies all parallel pairs of morphisms in $\mathcal{C}$. Let K be the coverage Q_+J on $\mathcal{P}$. Since Q is full and bijective on objects, it is not hard to see that the explicit description of Q_+J given in the proof above may be simplified in this case: a sieve R on U in $\mathcal{P}$ is K-covering iff the sieve $\{f \mid Q(f) \in R\}$ is J-covering. Now the composite $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{P}^{\mathrm{op}}, \mathbf{Set}]$ satisfies the hypotheses of A4.6.10, since its first factor is an inclusion and the second is hyperconnected by A4.6.9; hence $\mathbf{Sh}(\mathcal{P}, K)$ is the image of this composite in the hyperconnected–localic sense as well as the surjection–inclusion sense. Thus, by A4.6.2(e), $\mathbf{Sh}(\mathcal{P}, K)$ is also the image in the hyperconnected–localic sense of the unique geometric morphism $\mathbf{Sh}(\mathcal{C}, J) \to \mathbf{Set}$.

Remark 2.3.21 In connection with Example 2.3.20, it now seems appropriate to extend the notion of 'ideal for a coverage', which we defined for posites in 1.1.16(e), to arbitrary small sites. If $(\mathcal{C}, T)$ is a small site (where T is, for the moment, not necessarily a Grothendieck coverage), we define a subset I of ob $\mathcal{C}$ to be a T-*ideal* if it is a sieve and, whenever we are given an object U of $\mathcal{C}$ and a cover $(U_j \to U \mid j \in J)$ such that all the U_j lie in I, then necessarily $U \in I$. It is easy to see that the subterminal objects of $\mathbf{Sh}(\mathcal{C}, T)$ correspond bijectively to T-ideals in $\mathcal{C}$ (and the direct proof in 1.1.16(e) that the T-ideals form a frame still works without the assumption that $\mathcal{C}$ is a poset); and, in the context of 2.3.20, we have a bijection between J-ideals in $\mathcal{C}$ and Q_+J-ideals in its preorder reflection. This provides an alternative proof that the site constructed in 2.3.20 yields the localic reflection of $\mathbf{Sh}(\mathcal{C}, J)$.

There is an important special case in which the description of the localic reflection provided by 2.3.20 may be appreciably simplified.

Lemma 2.3.22 *Let $\mathcal{C}$ be a small regular category, and let J be a coverage on $\mathcal{C}$ which contains the regular coverage. Then the localic reflection of* $\mathbf{Sh}(\mathcal{C}, J)$ *may be identified with* $\mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), J')$ *where J' is the coverage induced on* $\mathrm{Sub}_{\mathcal{C}}(1)$ *by J, as in* 2.2.2, *or equivalently the coverage $I^{\bullet}J$ where I:* $\mathrm{Sub}_{\mathcal{C}}(1) \to \mathcal{C}$ *is the inclusion functor. Moreover, under this identification, the hyperconnected geometric morphism from* $\mathbf{Sh}(\mathcal{C}, J)$ *to its localic reflection is identified with that induced by I as in* 2.3.4.

Proof First we recall that the *support* σA of an object A of $\mathcal{C}$ is defined to be the image of the unique morphism $A \to 1$; we may regard σ as a functor $\mathcal{C} \to \mathrm{Sub}_{\mathcal{C}}(1)$, left adjoint to the inclusion (cf. A1.3.1(ii)). We note that, since J contains the regular coverage, the composite $I\sigma$ is cover-preserving; for if R is a J-covering sieve on an object A, and S denotes the sieve on σA generated by $\{\sigma f \mid f \in R\}$, then the pullback of S along $A \twoheadrightarrow \sigma A$ contains R, and so S is

covering by axiom (L). Hence σ is cover-preserving when we equip $\mathrm{Sub}_{\mathcal{C}}(1)$ with the induced coverage.

Let $(\mathcal{P}, K)$ be the site constructed in 2.3.20 for the localic reflection. Since the codomain of σ is a preorder, it factors uniquely through $Q\colon \mathcal{C} \to \mathcal{P}$, say by $\overline{\sigma}\colon \mathcal{P} \to \mathrm{Sub}_{\mathcal{C}}(1)$. From the description of K given in 2.3.20, it is clear that $\overline{\sigma}$ is cover-preserving. Moreover, the regularity of $\mathcal{C}$ implies that σ preserves finite products; it is easy to see that Q also preserves them, and hence $\overline{\sigma}$ is a cartesian functor (even though σ itself is not, in general!). So by 2.3.3 composition with $\overline{\sigma}$ yields a functor $\overline{\sigma}^*\colon \mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), J') \to \mathbf{Sh}(\mathcal{P}, K)$. But since J contains the regular coverage, we know that for any object A of $\mathcal{C}$ the singleton $\{A \twoheadrightarrow \sigma A\}$ generates a J-covering sieve; hence its image under Q generates a K-covering sieve, and so any K-sheaf F on $\mathcal{P}$ necessarily satisfies $F(QA) \cong F(Q\sigma A)$ for all A. It now follows easily that $\overline{\sigma}^*$ is full as well as faithful, and that it is essentially surjective on objects; so the geometric morphism induced by $\overline{\sigma}$ is an equivalence.

Finally, the inverse image functor $\mathbf{Sh}(\mathcal{P}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ was described in the proof of 2.3.18 as 'compose with Q, and then apply the associated sheaf functor'; hence, if we regard its domain as $\mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), J')$, the description becomes 'compose with σ, and then apply the associated sheaf functor'. So the direct image of the geometric morphism is given by composition with the right adjoint I of σ, as in 2.3.4. $\square$

If a cartesian functor $F\colon (\mathcal{D}, K) \to (\mathcal{C}, J)$ between small cartesian sites is both cover-preserving and cover-reflecting, then the functor $F^*\colon \mathbf{Sh}(\mathcal{C}, J) \to \mathbf{Sh}(\mathcal{D}, K)$ is both the direct image and the inverse image of geometric morphisms, which form an adjoint pair in $\mathfrak{Top}$. This occurs, of course, in the situation considered in 2.2.3, where F^* is an equivalence; but there are other examples of interest.

Example 2.3.23 Let $(\mathcal{C}, J)$ be the site considered in A2.1.11(d): that is, $\mathcal{C}$ is a small full subcategory of $\mathbf{Sp}$ closed under passage to open subspaces, and a sieve R on an object X is J-covering iff it contains an open covering of X in the traditional sense, i.e. a jointly-surjective family of open inclusions with codomain X. Let $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$. Then, for any object X of $\mathcal{C}$, we may identify the topos $\mathcal{E}/l(X)$ with $\mathbf{Sh}(\mathcal{C}/X, J_X)$ as defined in 2.2.17. Now we may identify $\mathcal{O}(X)$ with a full subcategory of $\mathcal{C}/X$ in the obvious way; the inclusion $\mathcal{O}(X) \to \mathcal{C}/X$ preserves finite limits, and it both preserves and reflects covers when $\mathcal{O}(X)$ is equipped with its canonical coverage. Thus we obtain geometric morphisms

$$\mathbf{Sh}(X) \underset{r}{\overset{i}{\rightleftarrows}} \mathcal{E}/l(X)$$

such that i^* and r_* are both the functor obtained by restricting sheaves along the inclusion functor $\mathcal{O}(X) \to \mathcal{C}/X$. In particular, r is right adjoint to i in $\mathfrak{Top}$, since r^* is left adjoint to i^*. We note also that i_* is full and faithful (i.e. i is an

inclusion), since it is obtained by right Kan extension along a full and faithful functor; hence r is a local geometric morphism in the sense of 1.5.6.

Similar examples may be constructed using the sites of A2.1.11(e) and (f): if X is a smooth manifold (respectively, an affine scheme of finite type over a field k) then $\mathcal{O}(X)$ embeds in $\mathbf{Mf}/X$ (resp. in the category of affine schemes over X) in a way which preserves and reflects covers; so we get an inclusion from $\mathbf{Sh}(X)$ to a slice of the topos of A2.1.11(e) (resp. (f)) which has a left adjoint in $\mathfrak{Top}$. We shall have more to say about these examples, and others like them, in Section C3.6 below.

Suggestions for further reading: Artin et al. [36], Mac Lane & Moerdijk [751].

C2.4 Internal sites and pullbacks

As with the notion of frame (cf. Section C1.6), the notion of (small) site can be 'internalized' in any topos $\mathcal{S}$: formally, a site is simply a model of a certain higher-order theory (cf. Section D4.1) and we can consider models of this theory in $\mathcal{S}$. Further, we may study the category of ($\mathcal{S}$-valued) sheaves on any such internal site $(\mathbb{C}, T)$ (for which our notation will be $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$); this category is a topos, and in fact a subtopos of the internal diagram category $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ (cf. B2.3.11), by the appropriate $\mathcal{S}$-indexed analogue of the arguments in Section C2.2. And the Giraud theorem (2.2.8) 'relativizes' to the assertion that every bounded geometric morphism with codomain $\mathcal{S}$ is equivalent to one of the form $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T) \to \mathcal{S}$ for some internal site $(\mathbb{C}, T)$ in $\mathcal{S}$; we proved this, in slightly different language, in B3.3.4.

We need to introduce some standard notation for internal sites. (We shall assume familiarity with the notation developed in Sections B2.3 and B2.5 for internal categories.) It is convenient to think of an internal coverage on an internal category $\mathbb{C}$ as being given by a span

$$C_0 \xleftarrow{\ b\ } T \xrightarrow{\ c\ } PC_1$$

where C_0 and C_1 are the object of objects and object of morphisms of $\mathbb{C}$; T is the 'object of all covers', b assigns to each cover the object which it covers (the *base* of the cover), and c assigns the set of morphisms which is the covering family. The usual definition of a coverage would demand that the pair (b, c) should be jointly monic, but it clearly does no harm if we omit this requirement, that is if we allow covers to 'occur with multiplicity' as members of T. Since this relaxation has the effect of simplifying some of the arguments relating to change of base, we shall henceforth allow it.

On the other hand, we do require the 'book-keeping' condition

$$(\forall t\colon T)(\forall a\colon C_1)((a \in c(t)) \Rightarrow (d_0(a) = b(t)))$$

which says that the codomains of all the members of a cover coincide with the base of the cover; we shall denote this condition by (B). In 'diagrammatic' terms, it says that if we form the pullback

$$\begin{tikzcd} W \arrow[r, "p"] \arrow[d, "q"] & \in_{C_1} \arrow[d, "n"] \\ T \arrow[r, "c"] & PC_1 \end{tikzcd}$$

(where $(n, e): \in_{C_1} \rightarrowtail PC_1 \times C_1$ is, as usual, a tabulation of the universal relation $PC_1 \looparrowright C_1$), then the square

$$\begin{tikzcd} W \arrow[r, "ep"] \arrow[d, "q"] & C_1 \arrow[d, "d_0"] \\ T \arrow[r, "b"] & C_0 \end{tikzcd}$$

commutes.

We shall also require our coverages to satisfy the diagrammatic form of the coverage axiom (C), of course; and in this section we shall generally require them to be sifted in the sense defined after 2.1.3 – we denote the latter condition by (S). In our present context, a sifted coverage is simply one for which the morphism c factors through the subobject $\mathrm{Sv}(\mathbb{C}) \rightarrowtail PC_1$ of sieves on $\mathbb{C}$, where a subobject $S \rightarrowtail C_1$ is called a sieve if it is closed under right multiplication, i.e. the composite

$$S \times_{C_0} C_1 \rightarrowtail C_1 \times_{C_0} C_1 \cong C_2 \xrightarrow{d_1} C_1$$

factors through $S \rightarrowtail C_1$. (Recall the convention of Section B2.5 regarding the use of the symbol $\times_{C_0}$.)

One reason for working with sifted coverages is that it makes the formulation of axiom (C) rather simpler: to say that (T, b, c) satisfies (C) means that, if we form the subobject $Q \rightarrowtail T \times_{C_0} C_1 \times_{C_0} T$ which consists of 'triples (t', a, t) such that $aa' \in t$ for all $a' \in t'$', then the composite

$$Q \rightarrowtail T \times_{C_0} C_1 \times_{C_0} T \xrightarrow{\pi_{23}} C_1 \times_{C_0} T$$

is an epimorphism. We may similarly formulate the axioms (M) and (L) of 2.1.8 in diagrammatic form; we leave this to the reader. Note that we shall *not* automatically assume in this section that our coverages satisfy axiom (L); the reason for this will become apparent in 2.4.9 below.

When dealing with coverages on an internal poset $\mathbb{C}$, we shall (as in 1.1.16(e) and elsewhere) tend to replace the span $C_0 \leftarrow T \rightarrow PC_1$ by one of the form $C_0 \leftarrow T \rightarrow PC_0$; that is, we shall 'identify the morphisms in a cover with their domains'. We again leave to the reader the task of reformulating the axioms (B), (C), (S), (M) and (L) in terms of coverages presented in this way.

We shall be concerned in this section mainly with 'change of base'; that is, with the effect on internal sites in $\mathcal{S}$ of a geometric morphism $f\colon \mathcal{T} \to \mathcal{S}$. Of course, given a site $(\mathbb{C}, T)$ in $\mathcal{S}$, we know from B3.3.6 that the pullback of $(\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T) \to \mathcal{S})$ along f (exists and) is a bounded $\mathcal{T}$-topos; indeed, from the proof of that result, we know that it can be represented as $\mathbf{Sh}_{\mathcal{T}}(f^*\mathbb{C}, f^{\#}T)$ for some coverage $f^{\#}T$ on $f^*\mathbb{C}$. The problem, therefore, is to identify $f^{\#}T$.

Before tackling the general question, we shall deal with the problem of change of base for locales. Clearly, given a geometric morphism $f\colon \mathcal{T} \to \mathcal{S}$ and a locale Y in $\mathcal{T}$, we may construct the hyperconnected–localic factorization of the composite

$$\mathbf{Sh}_{\mathcal{T}}(Y) \longrightarrow \mathcal{T} \xrightarrow{\ f\ } \mathcal{S}$$

in order to obtain a locale $f_!(Y)$ in $\mathcal{S}$. We can be more explicit: $\mathcal{O}(f_!(Y))$ is simply $f_*(\mathcal{O}(Y))$, since the latter is the direct image under the composite above of the subobject classifier of $\mathbf{Sh}_{\mathcal{T}}(Y)$. (Recall from B2.3.7 that f_* preserves completeness of internal posets; it also preserves Heyting algebras, since the latter can be defined equationally (A1.5.11), and thus induces a functor $\mathbf{Frm}(\mathcal{T}) \to \mathbf{Frm}(\mathcal{S})$, whose 'dual' we denote henceforth by $f_!\colon \mathbf{Loc}(\mathcal{T}) \to \mathbf{Loc}(\mathcal{S})$.)

Since pullbacks of localic morphisms (exist and) are localic, we also have a functor $f^*\colon \mathbf{Loc}(\mathcal{S}) \to \mathbf{Loc}(\mathcal{T})$ defined by saying that the diagram

$$\begin{tikzcd}
\mathbf{Sh}_{\mathcal{T}}(f^*(X)) \arrow[r] \arrow[d] & \mathbf{Sh}_{\mathcal{S}}(X) \arrow[d] \\
\mathcal{T} \arrow[r, "f"] & \mathcal{S}
\end{tikzcd}$$

should be a pullback for each locale X in $\mathcal{S}$. The universal property of pullbacks, and functoriality of the hyperconnected–localic factorization, tells us immediately that f^* is right adjoint to $f_!$; equivalently, the corresponding functor $f^{\#}$ on frames (defined by $f^{\#}(\mathcal{O}(X)) = \mathcal{O}(f^*(X))$) is *left* adjoint to f_*. However, $f^{\#}$ is *not* simply f^* applied to frames (indeed, the latter does not in general preserve completeness of internal posets), which is why we have adopted a different notation for it. (On the other hand, if X is a *discrete* internal locale, that is $\mathcal{O}(X)$ is a full power object PX, then $f^{\#}(\mathcal{O}(X))$ is simply $P(f^*X)$, since we

know that

$$\begin{tikzcd} \mathcal{T}/f^*X \arrow[r] \arrow[d] & \mathcal{S}/X \arrow[d] \\ \mathcal{T} \arrow[r, "f"] & \mathcal{S} \end{tikzcd}$$

is a pullback. Thus our use of the notation f^* for locales is consistent with its use for objects, considered as discrete locales.)

To obtain an explicit description of the functor $f^\#$, it is actually convenient to extend it to a larger category, namely that of internal complete semilattices in $\mathcal{S}$ (cf. Section C1.1). It is of course obvious that f_* extends to a functor $\mathbf{CjSLat}(\mathcal{T}) \to \mathbf{CjSLat}(\mathcal{S})$ (the fact that it preserves morphisms of complete semilattices follows from the fact that it preserves adjunctions), so the problem is to define a left adjoint to this functor. But in fact this is easy, because the categories of complete semilattices are monadic over the underlying toposes; so we can simply use A1.1.3(i) to lift the left adjoint of $f_*\colon \mathcal{T} \to \mathcal{S}$ to one at the level of complete semilattices. Explicitly, suppose we have a free presentation of a complete semilattice A in $\mathcal{S}$, that is a coequalizer diagram

$$PR \overset{a}{\underset{b}{\rightrightarrows}} PG \twoheadrightarrow A$$

in $\mathbf{CjSLat}(\mathcal{S})$. (Recall from the proof of 1.1.3 that free complete semilattices are simply power objects; the argument given there was constructive, and so valid in any topos.) Then, since we necessarily have $f^\#(PX) \cong P(f^*X)$ for any object X, and $f^\#$ must preserve coequalizers, there will be a coequalizer diagram $P(f^*R) \rightrightarrows P(f^*G) \twoheadrightarrow f^\#A$ in $\mathbf{CjSLat}(\mathcal{T})$. And the two morphisms $P(f^*R) \rightrightarrows P(f^*G)$ are easily described: since a and b correspond to morphisms $R \rightrightarrows PG$, or equivalently to a pair of relations $R \looparrowright G$, we simply apply f^* to these to obtain a pair of relations $f^*R \looparrowright f^*G$, and then convert them back into complete semilattice morphisms $P(f^*R) \rightrightarrows P(f^*G)$.

Lemma 2.4.1

(i) *The adjunction $(f^\# \dashv f_*)$ is enriched over* $\mathbf{CjSLat}(\mathcal{S})$, *i.e., for complete semilattices A, B in $\mathcal{S}, \mathcal{T}$ respectively, we have $[A, f_*B] \cong f_*[f^\#A, B]$, where the square brackets denote the 'internal homs' of* 1.1.5.

(ii) *$f^\#$ is a strong monoidal functor, that is, we have $f^\#(\Omega_{\mathcal{S}}) \cong \Omega_{\mathcal{T}}$ and $f^\#(A \otimes B) \cong f^\#A \otimes f^\#B$, compatibly with the coherence isomorphisms.*

(iii) *$f^\#$ restricts to a functor* $\mathbf{Frm}(\mathcal{S}) \to \mathbf{Frm}(\mathcal{T})$, *left adjoint to the corresponding restriction of f_*.*

Proof (i) First we note that the adjunction $(f^\# \dashv f_*)$ extends in an obvious way to an $\mathcal{S}$-indexed adjunction; hence it is enriched over $\mathcal{S}$, i.e. $[A, f_*B]$

and $f_*[f^\# A, B]$ are isomorphic as objects of $\mathcal{S}$. To see that this isomorphism is an isomorphism of internal posets in $\mathcal{S}$, we observe that the order-relation B_1 on B is itself a complete join-semilattice, and a parallel pair $u, v \colon A \rightrightarrows f_* B$ of complete join-homomorphisms satisfies $u \leq v$ iff (u, v) factors through $f_*(B_1) \rightarrowtail f_*(B \times B)$. Hence the partial order on $[A, f_* B]$ may be identified with $[A, f_* B_1]$; and similarly $f_*[f^\# A, B_1]$ is the partial order on $f_*[f^\# A, B]$. So $[A, f_* B]$ and $f_*[f^\# A, B]$ are isomorphic as internal posets, and hence as complete join-semilattices.

(ii) The identity $f^\#(\Omega_{\mathcal{S}}) \cong \Omega_{\mathcal{T}}$ is immediate from the fact that $\Omega_{\mathcal{S}}$ is the free complete semilattice on 1 (and f^* preserves 1). The preservation of binary tensor products follows immediately from (i), since for any complete semilattice C in $\mathcal{T}$ we have natural bijections

$$\begin{array}{rcl}
f^\#(A \otimes B) & \longrightarrow & C \\
\hline
A \otimes B & \longrightarrow & f_*(C) \\
\hline
A & \longrightarrow & [B, f_*(C)] \\
\hline
f^\# A & \longrightarrow & [f^\# B, C] \\
\hline
f^\# A \otimes f^\# B & \longrightarrow & C \quad .
\end{array}$$

Alternatively, we could deduce the result from the explicit description of $f^\#$ above, by showing that f^* preserves the presentation of $A \otimes B$ given in the proof of 1.1.5.

(iii) It follows immediately from (ii) that $f^\#$ lifts to a functor

$$\mathbf{CMon}(\mathbf{CjSLat}(\mathcal{S})) \to \mathbf{CMon}(\mathbf{CjSLat}(\mathcal{T})) \, .$$

The functor $f_* \colon \mathbf{CjSLat}(\mathcal{T}) \to \mathbf{CjSLat}(\mathcal{S})$ is not in general strongly monoidal, but as the right adjoint of a strong monoidal functor it is automatically lax monoidal, i.e. we have canonical maps $\Omega_{\mathcal{S}} \to f_*(\Omega_{\mathcal{T}})$ and $f_*(A) \otimes f_*(B) \to f_*(A \otimes B)$ satisfying the appropriate compatibility conditions; so f_* does lift to a functor between the categories of commutative monoids, and the adjunction $(f^\# \dashv f_*)$ also lifts to this level. But f_* also maps $\mathbf{Frm}(\mathcal{T})$ into $\mathbf{Frm}(\mathcal{S})$, as we observed earlier; and from the description of the coreflection functor $\mathbf{CMon}(\mathbf{CjSLat}(\mathcal{S})) \to \mathbf{Frm}(\mathcal{S})$ given in the proof of 1.1.10, it is easily seen that f_* commutes up to isomorphism with these coreflections. So by the dual of A1.1.3(ii) the left adjoint of f_* may be lifted to a functor $\mathbf{Frm}(\mathcal{S}) \to \mathbf{Frm}(\mathcal{T})$, as required. □

As in 2.2.4(*b*), by a *basis* for a locale X (or for the corresponding frame $\mathcal{O}(X)$), we mean a subobject $B \rightarrowtail \mathcal{O}(X)$ which generates it as a complete join-semilattice, i.e. such that the induced **CjSLat** homomorphism $PB \to \mathcal{O}(X)$ is an epimorphism. We saw in 2.2.4(*b*) that if we regard B as a poset, with the ordering it inherits from $\mathcal{O}(X)$, and equip it with the coverage T such that $T(U)$ consists of all subsets of $\downarrow(U)$ whose join (in $\mathcal{O}(X)$) is U, then the T-ideals (in the sense of 1.1.16(*e*)) in B are exactly the subsets $\{U : B \mid U \leq V\}$ where

$V \in \mathcal{O}(X)$, and hence $\mathbf{Sh}(B,T) \simeq \mathbf{Sh}(X)$. But we can also express this by saying that

$$PC \rightrightarrows PB \twoheadrightarrow \mathcal{O}(X)$$

is a coequalizer in **CjSLat**, where C is the set of those subsets of B whose join in $\mathcal{O}(X)$ lies in B, and the two complete semilattice morphisms $PC \rightrightarrows PB$ are induced by the inclusion $C \rightarrowtail PB$ and the mapping $S \mapsto \{\bigvee S\}$.

The arguments in the preceding paragraph are constructive, and so make sense in an arbitrary topos. Hence we may obtain an explicit description of $f^{\#}(L)$, for a frame L in $\mathcal{S}$, in terms of a basis:

Lemma 2.4.2 *Let $f: \mathcal{T} \to \mathcal{S}$ be a geometric morphism; let L be an internal frame in $\mathcal{S}$, and let $m: B \rightarrowtail L$ be a basis for L. Form the pullback*

$$\begin{tikzcd} T \arrow[rr, "b"] \arrow[d, tail, "c"'] & & B \arrow[d, tail, "m"] \\ PB \arrow[r, "\exists m"] & PL \arrow[r, "\bigvee"] & L \end{tikzcd}$$

Then $f^{\#}L$ may be presented as the coequalizer in $\mathbf{CjSLat}(\mathcal{T})$ of the pair of complete-semilattice morphisms $P(f^*T) \rightrightarrows P(f^*B)$ which correspond to the morphisms

$$f^*T \xrightarrow{f^*c} f^*(PB) \xrightarrow{\phi_B} P(f^*B)$$

and

$$f^*T \xrightarrow{f^*b} f^*B \xrightarrow{\{\}} P(f^*B)\,.$$

(In particular, the coequalizer thus described is a frame.) □

Equivalently, $f^{\#}L$ may be presented by the internal site $(f^*B, f^{\#}T)$ in $\mathcal{T}$, where $f^{\#}T$ is the coverage given by the span

$$f^*B \xleftarrow{f^*b} f^*T \xrightarrow{\phi_B f^*c} P(f^*B)\,.$$

We note that, since ϕ_B is not a monomorphism in general (cf. 3.1.7 below), the pair $(f^*b, \phi_B f^*c)$ will not in general be monic even if (b, c) is. This is the reason why we chose to relax our definition of coverage at the beginning of this section; had we not done so, we should have had to replace f^*T by its image in $f^*B \times P(f^*B)$ at this point.

Note in particular that we may always take B to be the whole of L; in this case C is simply PL, but we note that the covers which we obtain in our internal site $(f^*L, f^{\#}PL)$ for $f^{\#}L$ are indexed by the object $f^*(PL)$ rather than $P(f^*L)$. Thus, unless the comparison map ϕ_L is an isomorphism, we will not

simply obtain the 'canonical coverage' on f^*L. This is another reflection of the fact that $f^\# L$ is not simply f^*L in general.

On the other hand, let us note a particular case of this construction where things do work rather simply. By a *coherent frame*, we mean one of the form IB, where B is a distributive lattice (cf. 1.1.3). For such a frame, we may take B itself (or rather the set of principal ideals) as a basis, as we noted in 2.2.4(b); and we may take the coverage on it to be indexed by the set KB of finite subsets of B. But K *does* commute with inverse image functors (cf. D5.4.12), so we deduce

Corollary 2.4.3 *If $L = IB$ is a coherent internal frame in $\mathcal{S}$, and $f\colon \mathcal{T} \to \mathcal{S}$ is any geometric morphism, then $f^\# L \cong I(f^*B)$.* $\square$

We now revert to arbitrary sites $(\mathbb{C}, T)$ in $\mathcal{S}$. The foregoing provides the motivation for considering the span

$$f^*C_0 \xleftarrow{f^*b} f^*T \xrightarrow{\phi_{C_1} f^*c} P(f^*C_1)$$

as our candidate for the coverage $f^\# T$ that we require to form the pullback

$$\begin{tikzcd} \mathbf{Sh}_{\mathcal{T}}(f^*\mathbb{C}, f^\# T) \arrow[r] \arrow[d] & \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T) \arrow[d] \\ \mathcal{T} \arrow[r, "f"] & \mathcal{S} \end{tikzcd} .$$

Lemma 2.4.4 *Let $\mathbb{C}$ be an internal category in a topos $\mathcal{S}$, and suppose given a span $(C_0 \leftarrow T \to PC_1)$. If T satisfies any of the conditions (B), (C), (S) or (M), then so does $f^\# T$ for any geometric morphism $f\colon \mathcal{T} \to \mathcal{S}$.*

Proof This is mostly straightforward verification, using the diagrammatic forms of the conditions which we described above and the fact that

$$\begin{tikzcd} f^*(\in_{C_1}) \arrow[r] \arrow[d, "f^*(n)"] & \in_{f^*C_1} \arrow[d, "n"] \\ f^*(PC_1) \arrow[r, "\phi_{C_1}"] & P(f^*C_1) \end{tikzcd}$$

is a pullback by the definition of ϕ_{C_1}. For example, it is easily seen that if a subobject $S \rightarrowtail C_1$ is a sieve, then f^*S is a sieve in $f^*\mathbb{C}$; hence there is a

commutative square

$$\begin{tikzcd} f^*(\mathrm{Sv}(\mathbb{C})) \arrow[r] \arrow[d, tail] & \mathrm{Sv}(f^*\mathbb{C}) \arrow[d, tail] \\ f^*(PC_1) \arrow[r, "\phi_{C_1}"] & P(f^*C_1) \end{tikzcd} ,$$

from which it follows that $f^\# T$ inherits siftedness from T. □

In contrast to 2.4.4, the 'local character' condition (L) is *not* preserved under change of base; we shall see a counterexample in 2.4.9 below.

In proving that $f^\# T$ is indeed the 'pullback coverage' that we require, the key ingredient is the following:

Lemma 2.4.5 *Let $f: \mathcal{T} \to \mathcal{S}$ be a geometric morphism, $(\mathbb{C}, T)$ an internal site in $\mathcal{S}$, and $\mathbb{E}$ a cocomplete $\mathcal{T}$-indexed category. Then a $\mathcal{T}$-indexed functor $f^*\mathbb{C} \to \mathbb{E}$ sends covers in $f^\# T$ to epimorphic families in $\mathbb{E}$ iff the corresponding $\mathcal{S}$-indexed functor $\mathbb{C} \to (f^*)^*(\mathbb{E})$ (cf.* B2.3.14*) sends covers in T to epimorphic families.*

Proof This is really just a matter of unravelling the definitions. Let

$$\begin{tikzcd} W \arrow[r, "p"] \arrow[d, "q"] & \in_{C_1} \arrow[d, "n"] \\ T \arrow[r, "c"] & PC_1 \end{tikzcd}$$

be a pullback in $\mathcal{S}$; then in the diagram

$$\begin{tikzcd} f^*W \arrow[r, "f^*p"] \arrow[d, "f^*q"] & f^*(\in_{C_1}) \arrow[r] \arrow[d, "f^*n"] & \in_{f^*C_1} \arrow[d, "n'"] \\ f^*T \arrow[r, "f^*c"] & f^*(PC_1) \arrow[r, "\phi_{C_1}"] & P(f^*C_1) \end{tikzcd}$$

both squares are pullbacks (the right-hand one by the definition of ϕ_{C_1}), i.e. f^*W is the pullback of $\in_{f^*C_1}$ along the morphism c' of the coverage $f^\# T$. Thus to say that $F: f^*\mathbb{C} \to \mathbb{E}$ sends covers in $f^\# T$ to epimorphic families is to say that, if we form the object $A = F(f^*b: f^*T \to f^*C_0)$ of the fibre $\mathcal{E}^{f^*T}$, and the morphism $F(e'p': f^*W \to f^*C_1): B \to (f^*q)^*(A)$ in $\mathcal{E}^{f^*W}$ (where p' is the top composite in the second diagram above), then the transpose $\Sigma_{f^*q}(B) \to A$ of the latter is

an epimorphism in $\mathcal{E}^{f^*T}$. But this condition involves only the fibres of $\mathbb{E}$ over objects of the form f^*I, so it is easily seen to be the same as the condition that F sends covers in T to epimorphic families when considered as a functor $\mathbb{C} \to (f^*)^*\mathbb{E}$. □

Theorem 2.4.6 *Let $(\mathbb{C}, T)$ be an internal site in a topos $\mathcal{S}$ (where the coverage T is assumed to be sifted, but not necessarily a Grothendieck coverage), and $f\colon \mathcal{T} \to \mathcal{S}$ a geometric morphism. Then there is a pullback square*

$$\begin{tikzcd}
\mathbf{Sh}_{\mathcal{T}}(f^*\mathbb{C}, f^{\#}T) \arrow[r] \arrow[d] & \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T) \arrow[d] \\
\mathcal{T} \arrow[r, "f"] & \mathcal{S}
\end{tikzcd}$$

in $\mathfrak{Top}$.

Proof Let $g\colon \mathcal{E} \to \mathcal{T}$ be a $\mathcal{T}$-topos. We shall show that geometric morphisms $\mathcal{E} \to \mathbf{Sh}_{\mathcal{T}}(f^*\mathbb{C}, f^{\#}T)$ over $\mathcal{T}$ correspond to geometric morphisms $\mathcal{E} \to \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ over $\mathcal{S}$; this is clearly sufficient. But the latter correspond to $\mathcal{S}$-indexed functors $\mathbb{C} \to (f^*)^*(g^*)^*(\mathbb{E})$ (where $\mathbb{E}$ denotes the canonical indexing of $\mathcal{E}$ over itself) which are $\mathbb{C}^{\mathrm{op}}$-torsors in the sense of B3.2.3 and send covers in T to epimorphic families, by the $\mathcal{S}$-indexed version of 2.3.9. The passage between such functors and $\mathcal{T}$-indexed functors $f^*(\mathbb{C}) \to (g^*)^*(\mathbb{E})$ preserves and reflects the property of being a torsor, and we have seen in 2.4.5 that it preserves and reflects the condition on covers. So the result is established. □

Remark 2.4.7 By analogy with 2.4.5, one might hope to show that the correspondence of B2.3.14 between $\mathcal{T}$-indexed functors $f^*\mathbb{C}^{\mathrm{op}} \to \mathbb{T}$ and $\mathcal{S}$-indexed functors $\mathbb{C}^{\mathrm{op}} \to (f^*)^*(\mathbb{T})$ identifies the $f^{\#}T$-sheaves on $f^*\mathbb{C}$ with those functors $\mathbb{C}^{\mathrm{op}} \to (f^*)^*(\mathbb{T})$ which satisfy the sheaf axiom for covers in T. Unfortunately, this is not true in general. The reason is that, for a contravariant $F\colon f^*\mathbb{C}^{\mathrm{op}} \to \mathbb{T}$, we obtain (in the notation of 2.4.5) a morphism $(f^*q)^*(A) \to B$ in $\mathcal{T}/f^*W$, which transposes to a morphism $A \to \Pi_{f^*q}(B)$ in $\mathcal{T}/f^*T$, which must be the equalizer of two other morphisms (also defined using Π-functors) if F is to be an $f^{\#}T$-sheaf. But f^* does not commute with Π-functors in general, so this is not equivalent to the condition for F (in its other incarnation as an $\mathcal{S}$-indexed functor) to be a T-sheaf. However, there are two particular cases in which we can make this identification: one is when the change-of-base morphism f is locally connected, as defined in 3.3.1 below, and therefore does commute with Π-functors. The other is when $\mathbb{C}$ is cartesian (or at least has pullbacks) and every cover in T is finite (that is, indexed by a finite cardinal); for in this case we only need to consider Π-functors along morphisms $q\colon W \to T$ which are finite cardinals in $\mathcal{S}/T$, and since inverse image functors preserve exponentiation to

the power of a finite cardinal (D5.2.11), they also commute with such Π-functors (cf. D5.2.13).

Thus, in either of these cases, we may interpret 2.4.6 as saying that the pullback of $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ along f is nothing other than the category of 'T-sheaves on $\mathbb{C}$ with values in $\mathbb{T}$'. Note in particular that the second case includes the case where $\mathbb{C}$ is a coherent internal category and T is its coherent coverage; thus we may view this result as the generalization of 2.4.3 above to the non-localic case.

Next, we establish a criterion in terms of internal sites for surjectivity of bounded geometric morphisms, which we shall use frequently in the next chapter.

Lemma 2.4.8 *Let $(\mathbb{C}, J)$ be an internal site in $\mathcal{S}$, such that $\mathbb{C}$ has a terminal object $t\colon 1 \to C_0$ and J is a Grothendieck coverage. Then the following are equivalent:*

(i) *The canonical geometric morphism $g\colon \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J) \to \mathcal{S}$ is a surjection.*

(ii) *'Every cover of t is inhabited'; that is, if we form the pullback*

$$\begin{tikzcd} Q \arrow[r] \arrow[d] & J \arrow[d, "b"] \\ 1 \arrow[r, "t"] & C_0 \end{tikzcd},$$

then the composite $Q \to J \to PC_1$ factors through the object P^+C_1 of inhabited subobjects of C_1, i.e. the image of $n\colon \in_{C_1} \to PC_1$.

Proof Since hyperconnected morphisms are surjective, the morphism of (i) is surjective iff its localic part is surjective; but by 1.5.1(ii) this happens iff the unique frame homomorphism $\lambda\colon \Omega_{\mathcal{S}} \to g_*(\Omega_{\mathbf{Sh}(\mathbb{C},J)})$ is monic. We recall from 2.3.21 that we may identify the codomain of this morphism with the poset of J-ideals of $\mathbb{C}$; its right adjoint λ_* sends a J-ideal $I \rightarrowtail C_0$ to the truth-value $[\![I = C_0]\!]$, or equivalently to $[\![t \in I]\!]$, and so $\lambda(p)$ must be the largest J-ideal I for which $(t \in I) \Rightarrow p$ is valid. Consider the set

$$I(p) = \{u\colon C_0 \mid (\exists j\colon J)((b(j) = u) \wedge (\sigma(j) \Rightarrow p))\}$$

where $\sigma(j)$ denotes the assertion that $c(j)$ is inhabited. Using the local character axiom, it is easily verified that $I(p)$ is a J-ideal, and using axiom (M) we see that if p holds then $I(p)$ is the whole of C_0, i.e. $p \Rightarrow \lambda_* I(p)$. On the other hand, for any J-ideal K we have

$$(\forall u\colon C_0)((u \in I(\lambda_*(K))) \Rightarrow (u \in K))$$

since the hypothesis implies that u has a cover all of whose members lie in K. So the mapping $p \mapsto I(p)$ is the required left adjoint to λ_*.

Thus $\mathbf{Sh}(\mathbb{C}, J) \to \mathcal{S}$ is surjective iff we have

$$(\forall p\colon \Omega)((t \in I(p)) \Rightarrow p) .$$

But this is exactly the statement that every cover of t is inhabited. $\square$

A site satisfying the hypotheses of 2.4.8 is called *consistent.* In particular, a frame $\mathcal{O}(X)$ (or the corresponding locale X) is called consistent if every covering of its top element is inhabited; since the canonical coverage on a frame clearly has local character, this is equivalent to saying that the geometric morphism $\mathbf{Sh}_{\mathcal{S}}(X) \to \mathcal{S}$ is surjective. (Note that, in a non-Boolean topos $\mathcal{S}$, it is strictly stronger than saying that $\mathcal{O}(X)$ is *nontrivial*, i.e. that its top and bottom elements are disjoint; the latter is equivalent to saying that $\mathbf{Sh}_{\mathcal{S}}(X) \to \mathcal{S}$ has dense image.) Note also that the local character condition (L) really is needed in the proof of 2.4.8: if we are given a general coverage T and we merely know that every cover of some particular object t is inhabited, this tells us nothing about 'covers of covers of t', and so the Grothendieck coverage generated by T (defined as in 2.1.9) may fail to have the corresponding property. (But compare 3.1.18 below.)

Example 2.4.9 We may now, as promised, give an example to show that condition (L) is not stable under change of base. For the condition 'every cover of t is inhabited' is clearly stable under change of base, as are all the other hypotheses of 2.4.8; hence if (L) were also stable, the class of bounded surjections would be stable under pullback in $\mathfrak{Top}$. But we know that this is not so; for we saw in 1.2.12 that epimorphisms are not stable under pullback in **Loc**, and by 1.5.1(ii) epimorphisms of locales correspond to surjections between localic toposes (and pullbacks of locales correspond to pullbacks of toposes, by 1.4.8).

In view of 2.4.9, we often need to replace the coverage $f^{\#}T$ obtained after a change of base by its 'Grothendieck completion' as in 2.1.9. In general, it is hard to give an explicit description of the latter; but in certain cases we can deduce properties of it from the fact that it has the same sheaves as the coverage from which we start. A particular example, which will be of use in Section C3.2, is the following:

Lemma 2.4.10 *Let $(\mathbb{C}, T)$ be an internal site in a topos $\mathcal{S}$, and U an object of $\mathbb{C}$ (that is, a morphism $1 \to C_0$ in $\mathcal{S}$). If U is T-irreducible (in the sense that 'the only T-covering sieve on U is the maximal sieve M_U' is true in the internal logic of $\mathcal{S}$), then it remains irreducible in the Grothendieck coverage $\tilde{T}$ generated by T.*

Proof If U is irreducible, then the ($\mathcal{S}$-indexed) functor $\mathbb{C}(U, -)\colon \mathbb{C} \to \mathbb{S}$ maps covers in T to epimorphic families in $\mathbb{S}$; hence, for any object A of $\mathcal{S}$, the functor $A^{\mathbb{C}(U,-)}\colon \mathbb{C}^{\mathrm{op}} \to \mathbb{S}$ is a T-sheaf. So by 2.1.9 it is also a $\tilde{T}$-sheaf; in particular, if

$(U_i \to U \mid i \in I)$ is a $\tilde{T}$-covering family, then the induced map

$$A^{\mathbb{C}(U,U)} \longrightarrow \prod_{i \in I} A^{\mathbb{C}(U,U_i)} \cong A^{\coprod_i \mathbb{C}(U,U_i)}$$

is monic. Putting $A = \Omega$ (so that $A^{(-)}: \mathcal{S}^{\mathrm{op}} \to \mathcal{S}$ is faithful by A2.2.3, and hence reflects monomorphisms), we deduce that $\coprod_{i \in I} \mathbb{C}(U, U_i) \to \mathbb{C}(U, U)$ is epic in $\mathcal{S}$. Pulling the latter back along the morphism $1 \to \mathbb{C}(U, U)$ which names the identity morphism on U, we deduce that (it is true in the internal logic of $\mathcal{S}$ that) some $U_i \to U$ must be split epic; equivalently, if the $U_i \to U$ form a sieve, then that sieve must be M_U. □

Our next aim in this section is to establish a Beck–Chevalley condition for internal frames in toposes. As we shall see, this is closely related to the pullback-stability of the hyperconnected–localic factorization of Section A4.6.

Theorem 2.4.11 *Suppose given a pullback square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

of toposes and geometric morphisms, and suppose either *that f is bounded* or *that g is bounded and $\mathcal{E}$ has a natural number object. Then*

(i) *The square*

$$\begin{tikzcd} \mathbf{Frm}(\mathcal{G}) \arrow[r, "k^{\#}"] \arrow[d, "g_*"] & \mathbf{Frm}(\mathcal{H}) \arrow[d, "h_*"] \\ \mathbf{Frm}(\mathcal{E}) \arrow[r, "f^{\#}"] & \mathbf{Frm}(\mathcal{F}) \end{tikzcd}$$

commutes; more precisely, the canonical natural transformation $f^{\#} g_* \to h_* k^{\#}$ is an isomorphism.

(ii) *If g is hyperconnected, so is h.*

Proof We note first that the two parts of the theorem are equivalent. If (ii) holds then the hyperconnected–localic factorization is stable under pullback, since we know that pullbacks of localic morphisms are localic. So, using the descriptions of $f^{\#}$ and g_* in terms of their effects on localic toposes, with which we began, we see that both ways round the square produce the same effect when applied to a localic topos over $\mathcal{G}$. On the other hand if (i) holds and g is

hyperconnected, then on chasing the subobject classifier of $\mathcal{G}$ around the Beck–Chevalley square we see that h_* must also preserve Ω, i.e. h is hyperconnected.

In the case when we assume f to be bounded, we have already proved (ii), in B3.3.7. So we need only consider the case when g is bounded, but f need not be; in this case we shall prove (i).

If g is bounded, then we may factor it as a composite

$$\mathcal{G} \xrightarrow{\ i\ } [\mathbb{C}^{\mathrm{op}}, \mathcal{E}] \longrightarrow \mathcal{E}$$

where $\mathbb{C}$ is an internal category in $\mathcal{E}$ (which may be assumed to be cartesian) and i is an inclusion; and it clearly suffices to prove the result for the two factors separately. In fact we shall prove it when g is of the form $[\mathbb{C}^{\mathrm{op}}, \mathcal{E}] \to \mathcal{E}$ and when g is localic; this also suffices, since every inclusion is localic (A4.6.2(a)). But in the case when g (and hence also h) is localic, there is nothing to prove; for in this case, the effect of g_* on localic toposes over $\mathcal{G}$ is simply that of composition with g, and similarly for h_*. So the assertion reduces, as in A2.2.4, to the fact that a composite of two pullback squares is a pullback.

We are thus reduced to showing that the square

$$\begin{tikzcd}
\mathbf{Frm}([\mathbb{C}^{\mathrm{op}}, \mathcal{E}]) \arrow[r, "{[\mathbb{C}^{\mathrm{op}}, f]^{\#}}"] \arrow[d] & \mathbf{Frm}([f^*\mathbb{C}^{\mathrm{op}}, \mathcal{F}]) \arrow[d] \\
\mathbf{Frm}(\mathcal{E}) \arrow[r, "f^{\#}"] & \mathbf{Frm}(\mathcal{F})
\end{tikzcd}$$

commutes, where $\mathbb{C}$ is an internal cartesian category in $\mathcal{E}$ and the vertical arrows are obtained by applying the direct image functors ('evaluate at 1', where 1 is the terminal object of the appropriate category) to internal frames. In fact we show that the diagram commutes when we replace **Frm** by **CjSLat** throughout; since we know that the functors on **Frm** are the restrictions of those on **CjSLat**, this will suffice. Now in 1.6.9 we gave a description of internal complete (semi)lattices in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ where $\mathcal{C}$ is a small cartesian category; the description was constructive, and so remains valid when **Set** is replaced by an arbitrary topos. That is, an object of $\mathbf{CjSLat}([\mathbb{C}^{\mathrm{op}}, \mathcal{E}])$ is just an $\mathcal{E}$-indexed functor from $\mathbb{C}^{\mathrm{op}}$ to the $\mathcal{E}$-indexed category $\mathbf{CLat}(\mathcal{E})$ of internal complete lattices in $\mathcal{E}$, satisfying the Beck–Chevalley condition for pullback squares in $\mathbb{C}$. Given such a functor, we may compose it with $f^{\#} \colon \mathbf{CLat}(\mathcal{E}) \to \mathbf{CLat}(\mathcal{F})$ – note that, although $f^{\#}$ was defined as a functor on complete semilattices, it in fact preserves morphisms of complete lattices since these are exactly the morphisms of **CjSLat** which have left adjoints in **CjSLat** – to obtain an $\mathcal{E}$-indexed functor $\mathbb{C}^{\mathrm{op}} \to (f^*)^*(\mathbf{CLat}(\mathcal{F}))$, which by B2.3.14 is essentially the same thing as an $\mathcal{F}$-indexed functor $f^*\mathbb{C}^{\mathrm{op}} \to \mathbf{CLat}(\mathcal{F})$. It is straightforward to verify that the Beck–Chevalley condition of 1.6.9 is preserved by this construction, and so we

have defined a functor $\mathbf{CjSLat}([\mathbb{C}^{\mathrm{op}}, \mathcal{E}]) \to \mathbf{CjSLat}([f^*\mathbb{C}^{\mathrm{op}}, \mathcal{F}])$. Moreover, it is easy to verify that this is left adjoint to the functor induced by f_*, so it is the functor $[\mathbb{C}^{\mathrm{op}}, f]^{\#}$ by uniqueness of adjoints. And since, as already indicated, the vertical arrows in the square above are given by evaluation at the terminal objects of $\mathbb{C}$ and of $f^*\mathbb{C}$, the commutativity of the square is now immediate. □

The assumption, in the case when g is bounded, that $\mathcal{E}$ has a natural number object is needed in order to justify restricting our attention to cartesian internal categories in the course of the proof: we need to be able to replace an arbitrary site of definition for $\mathcal{G}$ over $\mathcal{E}$ by one whose underlying category is cartesian, as in the proof of 2.2.8 (and cf. also B3.2.6). The same assumption is of course needed for the second case of the following corollary:

Corollary 2.4.12 *The hyperconnected–localic factorization of an arbitrary (resp. bounded) geometric morphism is stable under bounded (resp. arbitrary) pullback.* □

In fact the Beck–Chevalley square of 2.4.11(i) still commutes if we replace frames by complete join-semilattices throughout. We omit the details of the proof, which may be found in [560].

Another consequence of 2.4.11(ii) is worth stating explicitly:

Corollary 2.4.13 *For any topos $\mathcal{S}$, the localic reflection functor $\mathfrak{BTop}/\mathcal{S} \to \mathfrak{LTop}/\mathcal{S}$ (cf.* A4.6.12*) preserves finite products.*

Proof It clearly preserves the terminal object. Let $p_i \colon \mathcal{E}_i \to \mathcal{S}$ $(i = 1, 2)$ be bounded morphisms, and let $\mathcal{L}_i \to \mathcal{S}$ denote their localic reflections. Then we may form the diagram

$$\begin{tikzcd}
\mathcal{E}_1 \times_{\mathcal{S}} \mathcal{E}_2 \arrow[r] \arrow[d] & \mathcal{L}_1 \times_{\mathcal{S}} \mathcal{E}_2 \arrow[r] \arrow[d] & \mathcal{E}_2 \arrow[d] \\
\mathcal{E}_1 \times_{\mathcal{S}} \mathcal{L}_2 \arrow[r] \arrow[d] & \mathcal{L}_1 \times_{\mathcal{S}} \mathcal{L}_2 \arrow[r] \arrow[d] & \mathcal{L}_2 \arrow[d] \\
\mathcal{E}_1 \arrow[r] & \mathcal{L}_1 \arrow[r] & \mathcal{S}
\end{tikzcd}$$

in which all the squares are pullbacks; so 2.4.11(ii) implies that all four edges, and hence the diagonal, of the upper left square are hyperconnected. Similarly, the diagonal of the lower right square is localic; so $\mathcal{L}_1 \times_{\mathcal{S}} \mathcal{L}_2$ is the reflection of $\mathcal{E}_1 \times_{\mathcal{S}} \mathcal{E}_2$ in $\mathfrak{LTop}/\mathcal{S}$. □

We may now prove a long-promised result, namely the converse of B3.3.8(ii).

Proposition 2.4.14 *Let $p\colon \mathcal{E} \to \mathcal{S}$ be a bounded geometric morphism such that the diagonal $\Delta_p\colon \mathcal{E} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$ is an inclusion. Then p is localic.*

Proof Let A be an object of $\mathcal{E}$. Since $\mathcal{E}/A \times_{\mathcal{E}} \mathcal{E}/A \simeq \mathcal{E}/A \times A$ (cf. B3.2.14), we have a pullback square

$$\begin{tikzcd} \mathcal{E}/A \times A \arrow[r] \arrow[d] & \mathcal{E} \arrow[d, "\Delta_p"] \\ \mathcal{E}/A \times_{\mathcal{S}} \mathcal{E}/A \arrow[r] & \mathcal{E} \times_{\mathcal{S}} \mathcal{E} \end{tikzcd}$$

and hence the left vertical edge of this square is an inclusion. So by A4.6.12 and 2.4.13 its reflection in $\mathfrak{LTop}/\mathcal{S}$ is an inclusion into a product; equivalently, by 1.5.1(i) and 1.4.8, we have an inclusion $g\colon Y \to X \times X$ in $\mathbf{Loc}(\mathcal{S})$, where X and Y are the internal locales in $\mathcal{S}$ defined by $\mathcal{O}(X) = p_*(\Omega^A)$ and $\mathcal{O}(Y) = p_*(\Omega^{A \times A})$.

We now assume for notational simplicity that $\mathcal{S}$ is the classical topos of sets. Identifying $\mathcal{O}(X \times X)$ with the **CjSLat** tensor product $\mathcal{O}(X) \otimes \mathcal{O}(X)$, as in 1.1.11, we see that g^* sends a typical element $W = \bigvee_{i \in I}(U_i \otimes V_i)$ of $\mathcal{O}(X) \otimes \mathcal{O}(X)$ to $\bigcup_{i \in I}(U_i \times V_i)$. In particular, taking W to be the element $g_*(\Delta)$, where Δ is the diagonal subobject $A \rightarrowtail A \times A$ regarded as an element of $\mathcal{O}(Y)$, we have an expression for $\Delta = g^* g_*(\Delta)$ as a union of 'rectangles' $U_i \times V_i$. Applying the inverse image of the continuous map $X \to Y$ corresponding to Δ, we obtain $X = \bigcup_{i \in I}(U_i \cap V_i)$; but the fact that $U_i \times V_i \leq \Delta$ in $\mathcal{O}(Y)$ clearly implies that $U_i \cap V_i$ is a subterminal object. So we have expressed X as an I-indexed union of open sublocales which are subterminal objects; equivalently, we have constructed a diagram of the form

$$\begin{tikzcd} B \arrow[r, two heads] \arrow[d, tail] & A \\ p^*(I) & \end{tikzcd}$$

in $\mathcal{E}$. But we can do this for any object A of $\mathcal{E}$; so p is localic. □

For an alternative proof of 2.4.14 using descent theory, see 5.3.10 below. Next, we note one further consequence of 2.4.11:

Corollary 2.4.15 *A geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ is hyperconnected iff the functor $f^{\#}\colon \mathbf{Frm}(\mathcal{E}) \to \mathbf{Frm}(\mathcal{F})$ is full and faithful.*

Proof First suppose f is hyperconnected. Then for any localic $\mathcal{E}$-topos $g\colon \mathcal{E}' \to \mathcal{E}$, the top and right edges of the pullback square

$$\begin{tikzcd} \mathcal{F}' \arrow[r] \arrow[d] & \mathcal{E}' \arrow[d] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

form the hyperconnected–localic factorization of its diagonal; so the counit of the adjunction $(f_! \dashv f^*)$ between categories of internal locales, or equivalently the unit of the adjunction $(f^\# \dashv f_*)$ between categories of internal frames, is an isomorphism. Hence $f^\#$ is full and faithful. The converse follows from considering the component of the unit at the initial object $\Omega_{\mathcal{E}}$ of $\mathbf{Frm}(\mathcal{E})$. □

In A4.1.16, we introduced the Beck–Chevalley condition for *objects* of toposes (in what was then merely assumed to be a commutative square of geometric morphisms – though in practice the squares for which we are able to verify the condition are usually pullbacks). Of course, objects may be regarded as discrete internal locales; but, although functors of the form $f^\#$ preserve discrete locales, as we noted earlier, those of the form f_* usually do not. We cannot therefore hope to deduce a Beck–Chevalley condition for objects from the result of 2.4.11; and indeed it does not hold in general.

We need to introduce some terminology to describe those morphisms for which it does hold:

Definition 2.4.16 Let $f\colon \mathcal{F} \to \mathcal{E}$ be a geometric morphism.

(i) We say f is a *right (weak) Beck–Chevalley morphism* if, for every bounded morphism $g\colon \mathcal{G} \to \mathcal{E}$ with the same codomain, the pullback square

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

satisfies the (weak) Beck–Chevalley condition as defined in A4.1.16, i.e. the canonical natural transformation $\theta\colon g^* f_* \to k_* h^*$ is an isomorphism (resp. monic).

(ii) We say f is a *left (weak) Beck–Chevalley morphism* if, for every bounded g as above, the pullback square satisfies the 'opposite-handed' Beck–Chevalley condition that $f^* g_* \to h_* k^*$ is an isomorphism (resp. monic).

(iii) We say f is a *stable left/right (weak) Beck–Chevalley morphism* if the pullback of f along any bounded morphism is a left/right (weak) Beck–Chevalley morphism.

The assignment of the adjectives 'right' and 'left' to the first two parts of the above definition is more or less arbitrary. The convention that we have adopted is justified by the fact that, if f is a right Beck–Chevalley morphism, then it is the right adjoint f_* which appears as one of the functors involved in the Beck–Chevalley condition. We shall often abbreviate the properties defined in 2.4.16 by their initials; thus 'SLBC morphism' means 'stable left Beck–Chevalley morphism', 'RWBC morphism' means 'right weak Beck–Chevalley morphism' and so on. Fortunately, we shall see in the next chapter that the stable properties, at least, turn out to be equivalent to other properties that have less cumbersome names. (It would be possible to introduce yet another variable into 2.4.16, besides the choices of left or right, weak or strong Beck–Chevalley condition, and stable or 'unstable': we could also allow the variant where the morphism f is assumed to be bounded, and we consider its pullbacks along arbitrary geometric morphisms rather than bounded ones. However, it seems not to be worth making this distinction in practice.)

We note that A4.6.8 has an immediate reinterpretation in our new terminology:

Lemma 2.4.17 *Hyperconnected morphisms are both SLWBC and SRWBC.*

Proof By 2.4.11(ii), any pullback of a pullback of a hyperconnected morphism yields a square in which two opposite sides are hyperconnected. So by A4.6.8 the square satisfies both weak Beck–Chevalley conditions. □

We shall have more to say about Beck–Chevalley morphisms at several points in Chapter C3.

Suggestions for further reading: Joyal & Tierney [560], Kock & Schmidt [648].

C2.5 Fibrations of sites

We saw in 2.3.10 that any geometric morphism between Grothendieck toposes may be induced by a morphism of sites. However, this result is not best possible: there is a very special type of morphism of sites, called a fibration of sites, which suffices to obtain all such geometric morphisms, and such fibrations may be used to give characterizations of particular classes of geometric morphisms. The construction of a fibration of sites from a given geometric morphism involves consideration of internal sites, which is why we postponed it until after Section C2.4; however, in this section we shall not be concerned with change of base (except in the proof of 2.5.14 below), and so we shall assume for notational convenience that our base topos is **Set**. (Nevertheless, all our arguments are constructive,

and so they apply to morphisms between bounded $\mathcal{S}$-toposes, for any $\mathcal{S}$.) We shall also reinstate our standing assumption from Sections C2.2 and C2.3 that all coverages are Grothendieck coverages unless stated otherwise; however, we shall not assume as in Section C2.3 that the underlying categories of our sites are cartesian. (For the reason why not, see 2.5.9 below.)

In fact a fibration of sites, like a geometric morphism of toposes, is not a single functor but an adjoint pair of functors. The key to understanding the notion is the following result, which could equally well have been proved in Section C2.3:

Lemma 2.5.1 *Let $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ be small sites, and let $F: \mathcal{C} \to \mathcal{D}$ and $L: \mathcal{D} \to \mathcal{C}$ be functors with $(L \dashv F)$. Then F is a morphism of sites iff L is cover-reflecting.*

Proof As F has a left adjoint, it is automatically flat in the sense of 2.3.7, since the comma category $(V \downarrow F)$ has an initial object for each V; so we have to prove that F preserves covers iff L reflects them. First suppose F preserves covers; let $V \in \mathrm{ob}\,\mathcal{D}$ and let $R \in J(LV)$. Then the morphisms $F(f)$, $f \in R$, generate a K-covering sieve on FLV; pulling this back along the unit $\alpha_V: V \to FLV$ of the adjunction, we obtain a K-covering sieve S such that, for every $g \in S$, the composite $\alpha_V g$ factors through some $F(f)$, $f \in R$. But this is equivalent to saying that $L(g)$ factors through f, and so implies that $L(g) \in R$. Thus we have verified that L is cover-reflecting.

Conversely, suppose L reflects covers. Let $U \in \mathrm{ob}\,\mathcal{C}$, and let $R \in J(U)$. Pulling back along the counit $\beta_U: LFU \to U$, we obtain a J-covering sieve $\beta_U^*(R)$ on LFU; applying the definition, we deduce that the sieve

$$S = \{g: V \to FU \mid \beta_U L(g) \in R\}$$

is in $K(FU)$. But $\beta_U L(g)$ is simply the transpose $\overline{g}$ of g across the adjunction $(L \dashv F)$; and g itself factors as $F(\overline{g})\alpha_V$. Thus S is contained in the sieve generated by the images under F of members of R, and in particular the latter must be K-covering. $\square$

If the conditions of 2.5.1 are satisfied, then it is clear that the functor $L^*: [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is left adjoint to F^*, and hence that the left adjoint $F_!$ of $F^*: \mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ may be obtained as the composite aL^*i of 2.3.18(ii). In other words, we have

Scholium 2.5.2 *If the conditions of* 2.5.1 *are satisfied, then the geometric morphism* $\mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J)$ *induced by F as in* 2.3.4 *coincides with that induced by L as in* 2.3.18. $\square$

Now suppose we have a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between Grothendieck toposes, and suppose we have chosen a site $(\mathcal{C}, J)$ for $\mathcal{E}$ (where $\mathcal{C}$ is small, but not necessarily cartesian). By B3.1.10(ii), f is bounded, and so we can represent $\mathcal{F}$ as $\mathbf{Sh}_{\mathcal{E}}(\mathbb{D}, K)$ for some internal site $(\mathbb{D}, K)$ in $\mathcal{E}$. In Section C1.6 we saw how

to 'externalize' the site corresponding to an internal locale in a Grothendieck topos; here our immediate aim is to do something similar for an arbitrary site.

To begin with, let us consider the case where the coverages J and K are trivial; i.e. $\mathcal{E}$ is $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ and $\mathcal{F}$ is $[\mathbb{D}^{\mathrm{op}}, [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]]$. In this case we may regard $\mathbb{D}$ as a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Cat}$, and use the Grothendieck construction (cf. B1.3.5) to convert it into a split fibration $\mathcal{C} \rtimes \mathbb{D} \to \mathcal{C}$. (The use of the 'semidirect product' sign $\rtimes$ is inspired by the case of groups: if $\mathcal{C}$ is a group G in $\mathbf{Set}$, and $\mathbb{D}$ is a group H in $[G, \mathbf{Set}]$ (that is, a group on which G acts by automorphisms), then this is exactly the semidirect product $G \rtimes H$ as usually defined.) Thus objects of $\mathcal{C} \rtimes \mathbb{D}$ are pairs (U, V) with $U \in \mathrm{ob}\,\mathcal{C}$ and $V \in \mathrm{ob}\,\mathbb{D}(U)$, and morphisms $(U', V') \to (U, V)$ are pairs (a, b) where $a\colon U' \to U$ in $\mathcal{C}$ and $b\colon V' \to \mathbb{D}(a)(V)$ in $\mathbb{D}(U')$.

Lemma 2.5.3 *With the above notation,* $[\mathbb{D}^{\mathrm{op}}, [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]] \simeq [(\mathcal{C} \rtimes \mathbb{D})^{\mathrm{op}}, \mathbf{Set}]$.

Proof If F is a diagram of shape $\mathbb{D}^{\mathrm{op}}$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, represented by a discrete fibration $\gamma\colon \mathbb{F} \to \mathbb{D}$, we define

$$\tilde{F}(U, V) = \{x \in F_0(U) \mid \gamma_0(x) = V\},$$

with $\tilde{F}(a, b)(x) = F(b)(F_0(a)(x))$, where $F(b)$ denotes the appropriate cartesian morphism of $\mathbb{F}(U')$ lying over b. Conversely, given $G\colon (\mathcal{C} \rtimes \mathbb{D})^{\mathrm{op}} \to \mathbf{Set}$, we define $\overline{G_0}\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ by

$$\overline{G_0}(U) = \coprod_{V \in D_0(U)} G(U, V),$$

with $\overline{G_0}(a)(x) = G(a, 1)(x) \in G(U', D_0(b)(V))$. It is easy to see that $\overline{G_0}$ comes equipped with a morphism $g_0\colon \overline{G_0} \to D_0$ sending each element of $G(U, V)$ to V, and that the latter has a right $\mathbb{D}$-action, so that we obtain a discrete fibration $\overline{\mathbb{G}} \to \mathbb{D}$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, or equivalently a diagram $\overline{G}$ of shape $\mathbb{D}^{\mathrm{op}}$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. Further, the constructions $F \mapsto \tilde{F}$ and $G \mapsto \overline{G}$ are inverse to each other up to isomorphism. □

Note that the particular case of 2.5.3 in which $\mathbb{D}$ is a discrete internal category in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ (i.e. takes discrete categories as values) is just A1.1.7.

Next, we 'put the coverages back'. We claim that, given coverages J and K on $\mathcal{C}$ and $\mathbb{D}$ respectively, there exists a coverage $J \rtimes K$ on $\mathcal{C} \rtimes \mathbb{D}$ such that, under the above correspondence, $\tilde{F}$ is a $(J \rtimes K)$-sheaf iff F takes values in the category of J-sheaves and is itself a K-sheaf. Explicitly, we say that a sieve R on an object (U, V) is $(J \rtimes K)$-covering iff there exists an element S of $K(U)$ with $b(S) = V$ (in the notation developed in Section C2.4) such that if we regard the element $c(S)$ of $PD_1(U)$ as a subsheaf of $\mathcal{C}(-, U) \times D_1$, i.e. as a set of pairs (f, g) with $f\colon U' \to U$ in $\mathcal{C}$ and $g\colon V' \to \mathbb{D}(f)(V)$ in $\mathbb{D}(U')$, then for each such pair there is a J-covering sieve of morphisms $f'\colon U'' \to U'$ for which the pairs $(ff', \mathbb{D}(f')(g))$ all belong to R.

It is tedious but straightforward to verify that this is indeed a Grothendieck coverage on $\mathcal{C} \rtimes \mathbb{D}$, and that $\tilde{F}$ is a $(J \rtimes K)$-sheaf iff F is a K-sheaf with values in $\mathbf{Sh}(\mathcal{C}, J)$. Thus we have proved most of

Proposition 2.5.4 *Let $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$ be a Grothendieck topos, and $(\mathbb{D}, K)$ an internal site in $\mathcal{E}$. Then $\mathbf{Sh}_{\mathcal{E}}(\mathbb{D}, K) \simeq \mathbf{Sh}(\mathcal{C} \rtimes \mathbb{D}, J \rtimes K)$. Moreover, the projection functor $P\colon \mathcal{C} \rtimes \mathbb{D} \to \mathcal{C}$ is cover-reflecting, and induces the geometric morphism $\mathbf{Sh}_{\mathcal{E}}(\mathbb{D}, K) \to \mathcal{E}$ as in 2.3.18.*

Proof It remains to establish the second assertion of the proposition. But if R is any J-covering sieve on U, and V any object of $\mathbb{D}(U)$, then the sieve of all morphisms $(f, g)\colon (U', V') \to (U, V)$ for which $f\colon U' \to U \in R$ is $(J \rtimes K)$-covering (take S to be the top element of $K(U)$ in the definition); so P is cover-reflecting, and by 2.3.18 we know that the right Kan extension functor $\varprojlim_P$ maps sheaves to sheaves. Now, to compute $\varprojlim_P(G)(U)$, we have to form the limit of the diagram of shape $(P \downarrow U)^{\mathrm{op}}$ sending $((U', V), a\colon U' \to U)$ to $G(U, V)$; but this is precisely the limit in $\mathcal{S}$ which we need to compute to obtain $f_*(\overline{G})(V)$, where f_* is the direct image of the given morphism, i.e. the restriction of the functor $\varprojlim_{\mathbb{D}}$ to sheaves. So we have identified f with the geometric morphism induced by P. $\square$

Corollary 2.5.5 *Any geometric morphism between Grothendieck toposes may be induced by a cover-reflecting functor between sites.* $\square$

So far, we have not made any assumptions about the categories $\mathcal{C}$ and $\mathbb{D}$. But we may clearly assume that $\mathbb{D}$ has a terminal object, i.e. that each $\mathbb{D}(U)$ has a terminal object t_U and that the transition functors $\mathbb{D}(U) \to \mathbb{D}(U')$ induced by morphisms $U' \to U$ of $\mathcal{C}$ preserve terminal objects. Then we have a functor $T\colon \mathcal{C} \to \mathcal{C} \rtimes \mathbb{D}$ sending U to (U, t_U); and it is clear that T is a full and faithful right adjoint to the projection P. This prompts the following definition:

Definition 2.5.6 Let $(\mathcal{C}, J)$ and $(\mathcal{D}, K)$ be small sites. By a *fibration of sites* $(P, T)\colon (\mathcal{D}, K) \to (\mathcal{C}, J)$, we mean a pair of functors $P\colon \mathcal{D} \to \mathcal{C}$ and $T\colon \mathcal{C} \to \mathcal{D}$ such that P is left adjoint to T, T is cover-preserving (equivalently, by 2.5.1, P is cover-reflecting), T is full and faithful (equivalently, the counit $PT \to 1$ is an isomorphism), and P is a fibration in the sense of B1.3.3.

A fibration of sites is called a continuous fibration by some authors (cf. [828]), but we have chosen to avoid that name for reasons already explained in Section C2.3. As with geometric morphisms (cf. A4.1.1), there is an arbitrary choice of the direction in which we consider a fibration of sites to be pointing: we have chosen to make it point in the same direction as the geometric morphism which it induces – though it should be noted that this is the direction of the left adjoint P, whereas a geometric morphism f points in the direction of the right adjoint f_*.

We may now strengthen 2.5.5, as follows.

Corollary 2.5.7 *Any geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between Grothendieck toposes may be induced (as in 2.5.2) by a fibration of sites $(P, T): (\mathcal{D}, K) \to (\mathcal{C}, J)$. Moreover, the categories $\mathcal{C}$ and $\mathcal{D}$ may be chosen to be cartesian, and P to be a cartesian functor.*

Proof First we choose $(\mathcal{C}, J)$ to be a site of definition for $\mathcal{E}$; then we choose $(\mathbb{D}, \tilde{K})$ similarly for $\mathcal{F}$ as a bounded $\mathcal{E}$-topos, but additionally requiring $\mathbb{D}$ to contain the terminal object of $\mathcal{F}$. Then we define $\mathcal{D} = \mathcal{C} \rtimes \mathbb{D}$ as in 2.5.3, and equip it with the coverage derived from those on $\mathcal{C}$ and $\mathbb{D}$ as in 2.5.4. The arguments above show that the projection $\mathcal{D} \to \mathcal{C}$ and its right adjoint have the required properties. For the final assertion, we may choose both $\mathcal{C}$ and $\mathbb{D}$ to be cartesian (i.e. closed under finite limits in their respective toposes): for $\mathbb{D}$, this means that the corresponding functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Cat}$ takes cartesian categories and functors as values. So it follows from B1.4.1 that the Grothendieck category $\mathcal{C} \rtimes \mathbb{D}$ has finite limits, and P preserves them. □

Note that we have actually proved rather more than we claimed in the statement of 2.5.7: the 'base site' $(\mathcal{C}, J)$ of the fibration may be chosen to be any site of definition for $\mathcal{E}$ over **Set**, although the choice of $(\mathcal{D}, K)$ will in general depend on f as well as $\mathcal{F}$.

Next, we look at some particular examples of geometric morphisms from the point of view of the results established in this section.

Examples 2.5.8 (*a*) Let $f: [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ be the geometric morphism induced as in A4.1.4 by a functor $F: \mathcal{D} \to \mathcal{C}$. We may factor F as

$$\mathcal{D} \xrightarrow{R} \mathbf{Gl}(F) \xrightarrow{P} \mathcal{C}$$

where $\mathbf{Gl}(F)$, as in A2.1.12, denotes the comma category $(\mathcal{C} \downarrow F)$, P is the projection $(U, V, \alpha) \mapsto U$ and R is the functor $V \mapsto (FV, V, 1_{FV})$ (the right adjoint of the other projection). P is a split fibration, corresponding to the functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Cat}$ which sends U to $(U \downarrow F)$, and a morphism $a: U' \to U$ to the functor induced by composition with a; R is full and faithful, and so if we use it to equip $\mathbf{Gl}(F)$ with a coverage K as in 2.2.4(*d*), we obtain an equivalence $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \simeq \mathbf{Sh}(\mathbf{Gl}(F), K)$. It is clear that P is cover-reflecting, since there are no nontrivial covers on $\mathcal{C}$; for an arbitrary F, P will not have a right adjoint, but if we assume that $\mathcal{D}$ has a terminal object and F preserves it, then we have a right adjoint T sending U to $(U, 1, U \to 1)$. So in this case we may take the fibration of sites $(P, T): (\mathbf{Gl}(F), K) \to (\mathcal{C}, D)$ (where D is the 'discrete' coverage on $\mathcal{C}$, whose only covers are the maximal sieves) to represent the geometric morphism f.

(*b*) Let $f: Y \to X$ be a locale morphism. We saw in 1.6.1 how to 'internalize' $\mathcal{O}(Y)$ as a frame in $\mathbf{Sh}(X)$; if we then externalize it again (not by the method of

1.6.2 but) by the method of 2.5.4, we obtain the posite $\mathcal{P}$ whose objects are pairs $(U,V) \in \mathcal{O}(X) \times \mathcal{O}(Y)$ with $V \leq f^*(U)$, with the coverage K in which a sieve R covers (U,V) iff the set of those V' for which there exists $U' \in \mathcal{O}(X)$ with $(U',V') \in R$ covers V in $\mathcal{O}(Y)$. In particular, we note that the single morphism $(U,V) \to (X,V)$ generates a covering sieve for any $V \leq f^*(U)$, and hence any sheaf A on $\mathcal{P}$ satisfies $A(U,V) \cong A(X,V)$, i.e. it is determined by its values on objects of the form (X,V). But the latter must satisfy the sheaf axiom for covers of the form $((X,V_i) \to (X,V) \mid i \in I)$ where $V = \bigcup\{V_i \mid i \in I\}$; so we deduce that $\mathbf{Sh}(\mathcal{P},K)$ is equivalent to $\mathbf{Sh}(Y)$. The projection $P\colon \mathcal{P} \to \mathcal{O}(X)$ is cover-reflecting, since if $U = \bigcup\{U_i \mid i \in I\}$ then $((U_i, V \cap f^*(U_i)) \to (U,V) \mid i \in I)$ generates a K-covering sieve; also P is a fibration, corresponding to the functor $\mathcal{O}(X)^{\mathrm{op}} \to \mathbf{Poset}$ sending U to $\downarrow(f^*(U))$, and it has a right adjoint given by $U \mapsto (U, f^*(U))$. So we have represented the geometric morphism $\mathbf{Sh}(Y) \to \mathbf{Sh}(X)$ induced by f by a fibration of sites.

(*c*) More generally, let $(\mathcal{C}, J)$ be a small cartesian site, and let L be an internal frame in $\mathbf{Sh}(\mathcal{C}, J)$, considered as a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{CLat}$ satisfying the conditions of 1.6.10. Then we may externalize L by the method of 2.5.4, to obtain a category $\mathcal{C} \rtimes L$ whose objects are pairs (U,V) with $U \in \mathrm{ob}\,\mathcal{C}$ and $V \in L(U)$, with morphisms $(U',V') \to (U,V)$ taken to be morphisms $a\colon U' \to U$ of $\mathcal{C}$ for which $V' \leq L(a)(V)$ in $L(U')$ (or equivalently $\Sigma_a(V') \leq V$ in $L(U)$). The detailed description of the coverage in this case is left to the reader.

(*d*) A particular case which we have seen before: let G be a group, and let X be an internal locale in $[G, \mathbf{Set}]$. Since the forgetful functor $[G, \mathbf{Set}] \to \mathbf{Set}$ is logical (and hence preserves completeness of posets, by B2.3.9), an internal locale in $[G, \mathbf{Set}]$ is simply a locale in $\mathbf{Set}$ on which G acts by automorphisms. If we externalize the site in $[G, \mathbf{Set}]$ given by $\mathcal{O}(X)$ with its G-action and canonical coverage, we obtain the site $(\mathcal{O}_G(X), T)$ for the topos of G-equivariant sheaves on X, which we described in A2.1.11(*c*). The fibration $\mathcal{O}_G(X) \to G$ is the faithful functor arising from the description of $\mathcal{O}_G(X)$ as a category structured over G, which we gave in A2.1.11(*c*); its right adjoint sends the unique object of G to the top element X of $\mathcal{O}(X)$.

(*e*) A generalization of (*d*) which will be of importance in Section C5.4: let G be a topological group, and consider an internal frame A in the topos $\mathbf{Cont}(G)$ of continuous G-sets. Since the geometric morphism $[G, \mathbf{Set}] \to \mathbf{Cont}(G)$ whose inverse image is the forgetful functor is hyperconnected, we know from 2.4.15 that $\mathbf{Frm}(\mathbf{Cont}(G))$ is a coreflective subcategory of $\mathbf{Frm}([G, \mathbf{Set}])$; hence A is the 'continuous part' of a frame $\hat{A}$ in $[G, \mathbf{Set}]$, that is a frame in $\mathbf{Set}$ on which G acts (not necessarily continuously) by automorphisms. (Explicitly, $\hat{A}$ may be constructed as the *MacNeille completion* of A, that is the poset of pairs (L,U) of subsets of A such that L is the set of lower bounds of U, and U is the set of upper bounds of L; we recall that the MacNeille completion of any Heyting algebra is a Heyting algebra [520, III 3.11], and hence a frame, and any G-action on a poset extends in an obvious way to its MacNeille completion. However, we do not need this identification for our present purposes.)

We recall that in 2.2.4(c) we constructed a site of definition $(\mathcal{C}, J)$ for $\mathbf{Cont}(G)$: the objects of $\mathcal{C}$ are the open subgroups of G, with morphisms $H \to K$ labelled by cosets gK of K such that $H \subseteq gKg^{-1}$, and every inhabited sieve is J-covering. Now, given an internal frame A in $\mathbf{Cont}(G)$, let A_H denote the set of elements of A which are invariant under the action of H; note that $A_K \subseteq A_H$ if $H \subseteq K$, and that the union of the A_H is the whole of A. The information that A is the continuous part of a complete Heyting algebra in $[G, \mathbf{Set}]$ tells us that the inclusions $A_K \to A_H$ are Heyting algebra homomorphisms, and that they have left and right adjoints (the left adjoint sends an element U of A_H to the join of its images under all the elements of K). Applying the construction of 2.5.4, we obtain a site $(\mathcal{D}, J')$ where the objects of $\mathcal{D}$ are pairs (H, U) with H an open subgroup of G and $U \in A_H$, and whose morphisms $(H, U) \to (K, V)$ are labelled by left cosets gK of K such that $H \subseteq gKg^{-1}$ and $U \leq gV$ in A. We note that any morphism gK of $\mathcal{D}$ has a factorization

$$(H, U) \xrightarrow{Hg} (g^{-1}Hg, g^{-1}U) \xrightarrow{K} (K, V)$$

where the first factor is an isomorphism; so a sieve on (K, V) is determined by the morphisms in it which are labelled by K itself. Thus we may define a sieve R on (K, V) to be J'-covering if V is the join of all those U for which there exists H such that $U \in A_H$ and $K\colon (H, U) \to (K, V) \in R$. In particular, we note that the single morphism $K\colon (H, V) \to (K, V)$ generates a covering sieve for any $H \subseteq K$ and any $V \in A_K$; so the projection $P\colon \mathcal{D} \to \mathcal{C}$ is cover-reflecting. Its right adjoint T is of course given by $H \mapsto (H, \top)$, where $\top$ is the top element of A.

Remark 2.5.9 If $\mathbb{C}$ has and P preserves pullbacks, then the fact that P is a fibration follows from the other conditions of 2.5.6: we may obtain a prone lifting of an arbitrary morphism $V \to PU$ by applying T to it and then pulling back along the unit $U \to TPU$. However, we do not wish to assume the existence or preservation of pullbacks as part of our definition of a fibration of sites: in part this is because we wish to be able to use the concept over base toposes which do not have a natural number object, so that we cannot in general choose our sites to be closed under finite limits (and thus we have to sacrifice the final assertion of 2.5.7). But we also wish to use 2.5.7 as a means of 'externalizing' characterizations of classes of geometric morphisms in terms of sites, and the latter may not be 'stable' under the operation of adjoining finite limits to the underlying category of a site.

To explain what we mean by the preceding sentence, let P be a property of geometric morphisms, and suppose we have a property P′ of sites (with terminal objects) such that a bounded geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$ has property P iff $\mathcal{E}$ is definable by an internal site in $\mathcal{S}$ having property P′. Then we may use the ideas above to 'externalize' P′ to a property P″ of fibrations of sites, such that a morphism of bounded $\mathcal{S}$-toposes has property P iff it may be induced by

a fibration of sites having property P″. We give one example of this technique here; others will occur at various points in Chapter C3.

Proposition 2.5.10 *A morphism of* $\mathfrak{BTop}/\mathbf{Set}$ *is surjective iff it can be induced by a fibration of sites* (P, T) *such that* P *preserves covers of objects in the image of* T.

Proof By 2.4.8, $f: \mathcal{F} \to \mathcal{E}$ is surjective iff we can choose a site $(\mathbb{D}, K)$ for $\mathcal{F}$ over $\mathcal{E}$ such that $\mathbb{D}$ has a terminal object, and every covering of the terminal object is inhabited. If $\mathcal{E}$ is itself given as $\mathbf{Sh}(\mathcal{C}, J)$, then this condition readily translates into the one in the statement of the proposition for the corresponding fibration of sites $(\mathcal{C} \rtimes \mathbb{D}, J \rtimes K) \to (\mathcal{C}, J)$. Conversely, if f is induced by a fibration of sites $(P, T): (\mathcal{D}, K) \to (\mathcal{C}, J)$ satisfying the condition, then 2.3.4 and 2.3.18 tell us that the composite $f_* f^*$ sends a sheaf A on $\mathcal{C}$ to the functor $a(A \circ P) \circ T$, where a is the associated K-sheaf functor. But from the fact that P preserves covers of objects in the image of T, we see that if two elements of $AP(TU) \cong A(U)$ become equal when restricted to some covering sieve R on TU, then they become equal on the cover of $PTU \cong U$ generated by the Pb, $b \in R$, and hence they must already be equal; so, from the construction of the associated sheaf functor in the proof of 2.2.6, we see that the unit map $(A \circ P) \to a(A \circ P)$ is monic at each object of the form TV. Hence the unit of $(f^* \dashv f_*)$ is monic, i.e. f is a surjection. □

Another application of fibrations of sites occurs in the description of (cofiltered) inverse limits in $\mathfrak{BTop}/\mathbf{Set}$ (or more generally in $\mathfrak{BTop}/\mathcal{S}$, where $\mathcal{S}$ is any topos with a natural number object). For the moment, we shall restrict ourselves to limits of inverse sequences

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

of Grothendieck toposes and geometric morphisms; dealing with more general cofiltered limits requires descent theory, and so we shall postpone it to Section C5.1. By the remark immediately after the proof of 2.5.7, we may successively choose sites $(\mathcal{C}_n, J_n)$ for each $\mathcal{E}_n$, and fibrations of sites $(P_n, T_n): (\mathcal{C}_n, J_n) \to (\mathcal{C}_{n-1}, J_{n-1})$ inducing the morphisms f_n as in 2.5.7.

We define a category $\mathcal{C}_\infty$ as follows: its objects are sequences $(U_n \mid n \geq 0)$ of objects of $\mathcal{C}_n$ for each n, equipped with isomorphisms $\alpha_n: P_n(U_n) \to U_{n-1}$ for each n such that, for some m, the transpose $U_n \to T_n(U_{n-1})$ of α_n is also an isomorphism whenever $n \geq m$. A morphism $(U_n \mid n \geq 0) \to (V_n \mid n \geq 0)$ of $\mathcal{C}_\infty$ is simply a sequance of morphisms $a_n: U_n \to V_n$ commuting in the obvious sense with the α_n; note that we shall tend to suppress the names of the α_n when referring to an object of $\mathcal{C}_\infty$, as is our usual custom with coherence isomorphisms. ($\mathcal{C}_\infty$ could equivalently be described as the pseudo-colimit in $\mathfrak{Cat}$ of the filtered diagram formed by the $\mathcal{C}_n$ and T_n; and indeed 2.5.11 below is really a special case of 2.3.14, except that we are here working with not-necessarily-cartesian

sites. However, the left adjoints P_n allow us to embed this pseudo-colimit as a subcategory of the pseudo-limit of the P_n, and this description is more useful for calculations.)

For $m \leq n$, we write $P_n^m : \mathcal{C}_n \to \mathcal{C}_m$ (resp. $T_n^m : \mathcal{C}_m \to \mathcal{C}_n$) for the composite of the P_j (resp. T_j) with $m+1 \leq j \leq n$ (we take this to be the identity functor if $m = n$, of course). We have functors $P_\infty^n : \mathcal{C}_\infty \to \mathcal{C}_n$ defined by $P_\infty^n(U_m \mid m \geq 0) = U_n$, and $T_\infty^n : \mathcal{C}_n \to \mathcal{C}_\infty$ defined by $T_\infty^n(U) = (V_m \mid m \geq 0)$, where $V_m = P_m^n(U)$ if $m \leq n$, and $V_m = T_m^n(U)$ if $m > n$, with the obvious coherence isomorphisms. It is easy to see that T_∞^n is right adjoint to P_∞^n, and the counit of the adjunction is an isomorphism. Also, P_∞^n is a fibration: given an object $(U_m \mid m \geq 0)$ of $\mathcal{C}_\infty$ and a morphism $a : V \to U_n = P_\infty^n(U_m \mid m \geq 0)$ in $\mathcal{C}_n$, we obtain a prone lifting of a by taking its mth component to be $P_m^n(a) : P_m^n(V) \to P_m^n(U_n) \cong U_m$ if $m \leq n$, and the appropriate prone lifting of a to $\mathcal{C}_m$ if $m > n$.

We equip $\mathcal{C}_\infty$ with the smallest coverage J_∞ for which the functors T_∞^n are all cover-preserving, i.e. the join (in the lattice of Grothendieck coverages on $\mathcal{C}_\infty$) of the coverages $(T_\infty^n)_\bullet(J_n)$ as defined in 2.3.12. By 2.5.1, this coverage also makes each P_∞^n cover-reflecting. So we have fibrations of sites $(P_\infty^n, T_\infty^n) : (\mathcal{C}_\infty, J_\infty) \to (\mathcal{C}_n, J_n)$, and hence geometric morphisms $f_\infty^n : \mathcal{E}_\infty = \mathbf{Sh}(\mathcal{C}_\infty, J_\infty) \to \mathcal{E}_n$, for each n; it is clear that these form a pseudo-cone over the diagram formed by the f_n.

Theorem 2.5.11 *With the above notation, $\mathcal{E}_\infty$ is the limit of the diagram formed by the $\mathcal{E}_n$ and f_n in $\mathfrak{BTop}/\mathbf{Set}$.*

Proof Suppose given an arbitrary (pseudo-)cone $(g_n : \mathcal{F} \to \mathcal{E}_n \mid n \geq 0)$, where $\mathcal{F}$ is a Grothendieck topos. By 2.3.9, each g_n corresponds to a flat cover-preserving functor $G_n : \mathcal{C}_n \to \mathcal{F}$ (where $\mathcal{F}$ is equipped with its canonical coverage); and we have coherence isomorphisms $G_n \circ T_n \cong G_{n-1}$ for all $n \geq 1$. Given an object $(U_n \mid n \geq 0)$ of $\mathcal{C}_\infty$, we define $G_\infty(U_n \mid n \geq 0)$ to be the limit in $\mathcal{F}$ of the diagram formed by the objects $G_n(U_n)$, with transition maps

$$G_n(U_n) \xrightarrow{G_n(\overline{\alpha_n})} G_n T_n(U_{n-1}) \xrightarrow{\sim} G_{n-1}(U_{n-1});$$

note that, for a given object $(U_n \mid n \geq 0)$, these transition maps are isomorphisms for all sufficiently large n, so we could equivalently define $G_\infty(U_n \mid n \geq 0)$ to be simply $G_m(U_m)$ for some sufficiently large m. Hence also we have canonical isomorphisms $G_\infty \circ T_\infty^n \cong G_n$ for all n, which are compatible with the coherence isomorphisms between the G_n. It follows easily that G_∞ is cover-preserving; for each $G_\infty \circ T_\infty^n$ is cover-preserving, and hence we have $(T_\infty^n)_\bullet(J_n) \leq (G_\infty)^\bullet(C)$ as coverages on $\mathcal{C}_\infty$, where C is the canonical coverage on $\mathcal{F}$. It is also easy to see that G_∞ is a flat functor, and hence that it induces a geometric morphism $g_\infty : \mathcal{F} \to \mathcal{E}_\infty$. Moreover, the functor G_∞ (and hence the geometric morphism g_∞) is determined up to canonical isomorphism by the requirement that we have $G_\infty \circ T_\infty^n \cong G_n$ for all n, since any object $(U_n \mid n \geq 0)$ of $\mathcal{C}_\infty$ is isomorphic to

$T_\infty^m(U_m)$ for sufficiently large m. And the construction of g_∞ from $(g_n \mid n \geq 0)$ is clearly functorial on 2-cells of $\mathfrak{BTop}/\mathbf{Set}$, so we have an equivalence

$$\mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E}_\infty) \simeq \varprojlim_n \mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E}_n),$$

as required. □

We note that, if we wished to avoid the use of the large site $(\mathcal{F}, C)$ in the proof of 2.5.11, it would be possible to replace it by a small site $(\mathcal{D}, K)$, where $\mathcal{D}$ is any full subcategory of $\mathcal{F}$ containing a generating set and also all objects of the form $g_n^*(l_n(U))$, $U \in \text{ob}\,\mathcal{C}_n$ (where $l_n \colon \mathcal{C}_n \to \mathbf{Sh}(\mathcal{C}_n, J_n)$ is the canonical functor), and K is the induced coverage. Thus the proof of 2.5.11 may be carried out in $\mathfrak{BTop}/\mathcal{S}$, for any base topos $\mathcal{S}$ with a natural number object.

We note the following important application of 2.5.11:

Corollary 2.5.12 *The localic reflection functor* $\mathfrak{BTop}/\mathbf{Set} \to \mathfrak{LTop}/\mathbf{Set}$ *preserves limits of inverse sequences.*

Proof We saw in 2.3.20 that we may construct the localic reflection at the level of sites, by forming the preorder reflection of the underlying category and equipping it with an appropriate coverage. If we are given an inverse sequence of fibrations of sites $(P_n, T_n) \colon (\mathcal{C}_n, J_n) \to (\mathcal{C}_{n-1}, J_{n-1})$ as in the proof of 2.5.11, and form the preorder reflections $\mathcal{P}_n$ of the categories $\mathcal{C}_n$, then the P_n and T_n clearly induce functors between the latter, and it is easy to verify that these are again fibrations of sites. Moreover, the construction of the limit site $(\mathcal{C}_\infty, J_\infty)$ also commutes with the preorder reflection; that is, the preorder $\mathcal{P}_\infty$ constructed from the $\mathcal{P}_n$ is the preorder reflection of $\mathcal{C}_\infty$, and the coverage which it inherits as a quotient of $\mathcal{C}_\infty$ coincides with that which it acquires as the limit of the $\mathcal{P}_n$. So the result is immediate from 2.5.11. □

Remark 2.5.13 Combining 2.5.12 with 2.4.13, we see that the localic reflection functor preserves countable products (and, once we have extended 2.5.12 to more general cofiltered limits in 5.1.12 below, we may similarly deduce that it preserves arbitrary $\mathcal{S}$-indexed products, where $\mathcal{S}$ is our base topos). However, it does not preserve arbitrary limits: the following counterexample is due to I. Moerdijk [828].

Let $\mathcal{C}$ be the finite category represented by the diagram

$$V \overset{a}{\underset{b}{\rightrightarrows}} U$$

and let $\mathbb{D}$ be the functor $\mathcal{C}^{\text{op}} \to \mathbf{Poset}$ defined by taking $\mathbb{D}(U) = \mathbf{1}$, $\mathbb{D}(V) = \mathbf{2}$, $\mathbb{D}(a)(0) = 0$ and $\mathbb{D}(b)(0) = 1$. (Here we are using $\mathbf{n}$ to denote the n-element totally ordered set $\{0, 1, \ldots, n-1\}$.) Let $\mathbb{E}$ be defined similarly,

but with the rôles of a and b interchanged. By B3.2.14 we have a pullback diagram

$$\begin{array}{ccc} [(\mathbb{D}\times\mathbb{E})^{\mathrm{op}},[\mathcal{C}^{\mathrm{op}},\mathbf{Set}]] & \longrightarrow & [\mathbb{E}^{\mathrm{op}},[\mathcal{C}^{\mathrm{op}},\mathbf{Set}]] \\ \downarrow & & \downarrow \\ [\mathbb{D}^{\mathrm{op}},[\mathcal{C}^{\mathrm{op}},\mathbf{Set}]] & \longrightarrow & [\mathcal{C}^{\mathrm{op}},\mathbf{Set}] \end{array}$$

and by 2.5.3 we may rewrite this in the form

$$\begin{array}{ccc} [(\mathcal{C}\rtimes(\mathbb{D}\times\mathbb{E}))^{\mathrm{op}},\mathbf{Set}] & \longrightarrow & [(\mathcal{C}\rtimes\mathbb{E})^{\mathrm{op}},\mathbf{Set}] \\ \downarrow & & \downarrow \\ [(\mathcal{C}\rtimes\mathbb{D})^{\mathrm{op}},\mathbf{Set}] & \longrightarrow & [\mathcal{C}^{\mathrm{op}},\mathbf{Set}] \end{array}\,.$$

But it is easy to compute these finite semidirect products explicitly, and to verify that if we take their preorder reflections we get a diagram which is not a pullback in **Poset**, and hence cannot give rise to a pullback of localic **Set**-toposes.

Using 2.5.12, we may extend the result of 1.1.12 from locales to toposes, at least in the countable case.

Corollary 2.5.14 *Let*

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

be an inverse sequence of Grothendieck toposes, such that the transition morphisms f_n are all surjections. Then the legs $f^n_\infty\colon \mathcal{E}_\infty \to \mathcal{E}_n$ of the limit cone in $\mathfrak{BTop}/\mathbf{Set}$ are also surjections.

Proof Fix a particular value of n, and consider the sequence in $\mathfrak{BTop}/\mathcal{E}_n$ formed by the $\mathcal{E}_m$ for $m \geq n$. If we replace each such $\mathcal{E}_m$ by its reflection ($\mathcal{L}_m$, say) in $\mathfrak{LTop}/\mathcal{E}_n$, then since hyperconnected morphisms are surjective, it is clear that the transition morphisms $\mathcal{L}_{m+1} \to \mathcal{L}_m$ are still surjective. So by 1.1.12 (applied to internal locales in $\mathcal{E}_n$) the inverse limit $\mathcal{L}_\infty$ of the $\mathcal{L}_m$ maps surjectively to each $\mathcal{L}_m$, and in particular to $\mathcal{L}_n = \mathcal{E}_n$; but, by 2.5.12 applied to bounded $\mathcal{E}_n$-toposes, the comparison $\mathcal{E}_\infty \to \mathcal{L}_\infty$ is hyperconnected, and hence surjective. □

In principle, it should be possible to prove 2.5.14 directly from the characterization of surjections in 2.5.10. The problem with trying to do so is that our definition of the limit coverage J_∞ gives us an explicit description only of

a generating set of covering sieves, not of
that P^n_∞ preserves those covers of objects
the generating family, but we cannot concl
objects.

Suggestions for further reading: Artin

all of them; it is not hard to verify
in the image of T^n_∞ which belong to
ıde that it preserves all covers of such

et al. [36], Moerdijk [828].

C3

CLASSES OF GEOMETRIC MORPHISMS

C3.1 Open maps

Our aim in this chapter is to revisit some of the 'topologically inspired' classes of geometric morphisms which we introduced in Section C1.5, in the context of morphisms between localic toposes, and to investigate their properties in the more general context of morphisms between Grothendieck toposes. As in Section C2.4, we shall henceforth interpret 'Grothendieck topos' loosely as meaning any topos defined and bounded over a base topos $\mathcal{S}$ having a natural number object; but we shall often treat $\mathcal{S}$ notationally as if it were the classical category **Set** of sets, relying on the reader to translate our arguments as required into the language of $\mathcal{S}$-indexed categories as developed in Part B.

We begin with the study of open maps. Recall that in 1.5.4 we defined a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ to be *open* if its inverse image f^* is a Heyting functor. However, this definition is equivalent to another elementary one, which is more convenient for use in some contexts; so we shall begin this section by establishing the equivalence.

As we saw in Section A2.1, any cartesian functor $F: \mathcal{E} \to \mathcal{F}$ between toposes comes equipped with a family of comparison morphisms $\phi_A: F(PA) \to P(FA)$ for each object A of $\mathcal{E}$ (in fact they form a natural transformation between functors $\mathcal{E}^{\mathrm{op}} \to \mathcal{F}$); explicitly, ϕ_A is the name of the relation tabulated by $F(\in_A) \rightarrowtail F(PA \times A) \cong F(PA) \times FA$. We call F a logical functor if ϕ_A is an isomorphism for any A; we shall call it *sub-logical* if we merely know that ϕ_A is monic for each A. Similarly, we shall call F a *sub-cartesian-closed functor* if the comparison map $\theta_{A,B}: F(B^A) \to FB^{FA}$ (defined in Section A1.5) is monic for all objects A and B of $\mathcal{E}$.

Lemma 3.1.1 *A cartesian functor $F: \mathcal{E} \to \mathcal{F}$ between toposes is sub-logical iff it is sub-cartesian-closed and $\phi_1: F\Omega_{\mathcal{E}} \to \Omega_{\mathcal{F}}$ is monic.*

Proof If F is sub-logical, then clearly ϕ_1 is monic; and from the description of the exponential B^A as a subobject of $P(A \times B)$ given in A2.4.11, it is easy to

see that the diagram

$$\begin{tikzcd}
F(B^A) \arrow[r, "\theta_{A,B}"] \arrow[d, tail] & FB^{FA} \arrow[d, tail] \\
F(P(A \times B)) \arrow[r, "\phi_{A\times B}"] & P(FA \times FB)
\end{tikzcd}$$

commutes. Hence $\theta_{A,B}$ is monic if $\phi_{A\times B}$ is. Conversely, it is not hard to see that ϕ_A may be decomposed as the composite

$$F(\Omega^A) \xrightarrow{\theta_{A,\Omega}} F\Omega^{FA} \xrightarrow{\phi_1^{FA}} \Omega^{FA},$$

so it is monic if $\theta_{A,\Omega}$ and ϕ_1 are. □

Lemma 3.1.2 *Let $f\colon \mathcal{C} \to \mathcal{D}$ be a functor between small categories. Then the functor $f^*\colon [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ induced by composition with f is sub-logical iff f satisfies the following condition: given any morphism $b\colon fU \to V$ in $\mathcal{D}$ whose domain is in the image of f, there exist $a\colon U \to U'$, $r\colon fU' \to V$ and $i\colon V \to fU'$ such that $ri = 1_V$ and $ib = fa$.*

Proof Suppose first that this condition holds: let A be an object of $[\mathcal{D}, \mathbf{Set}]$ and U an object of $\mathcal{C}$. Then, by the construction of A1.5.5 and A1.6.6, $f^*(PA)(U) = PA(fU)$ is the set of subfunctors of $A \times \mathcal{D}(fU, -)$, and $(\phi_A)_U$ maps such a subfunctor R to the subfunctor S of $f^*A \times \mathcal{C}(U, -)$ given by

$$S(U') = \{(x, a) \mid (x, fa) \in R(fU')\} .$$

But if $b\colon fU \to V$ in $\mathcal{D}$ and $y \in A(V)$, then (defining a, r and i as in the condition) it is easily seen that $(y, b) \in R(V)$ iff $(Ai(y), a) \in S(U')$; so R is uniquely recoverable from S, i.e. $(\phi_A)_U$ is injective.

Conversely, suppose f^* is sub-logical. Given $b\colon fU \to V$, let A be the functor $\mathcal{D}(V, -)$, and consider the following two subfunctors of $A \times \mathcal{D}(fU, -)$ as elements of $f^*(PA)(U)$:

$$R_1(V') = \{(c\colon V \to V', d\colon fU \to V') \mid cb = d\},$$

$$R_2(V') = \{(c, d) \in R_1(V') \mid c \text{ is not split monic}\} .$$

Since $(1_V, b) \in R_1(V)$, we have $R_1 \neq R_2$, and so their images under $(\phi_A)_U$ must be distinct; but this is equivalent to saying that we can find $i\colon V \to fU'$ and $a\colon U \to U'$ such that $(i, fa) \in R_1(fU')$ and i is split monic, i.e. that the given condition holds. □

Although the second half of the above argument is clearly non-constructive, the first half is constructive, and so can be applied to a functor $f : \mathbb{C} \to \mathbb{D}$ between internal categories in any topos $\mathcal{S}$, to obtain a sufficient condition for the corresponding inverse image functor $[\mathbb{D}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ to be sub-logical.

Remark 3.1.3 Reverting to the case $\mathcal{S} = \mathbf{Set}$, a similar argument shows that $f^* : [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ is sub-cartesian-closed iff f satisfies the condition that, given $b : fU \to V$, we can find $a : U \to U'$, $r : fU' \to V$ and $i : V \to fU'$ satisfying $ri = 1_V$ and $r(fa) = b$. We omit the details, which may be found in [513]; but we note that this result enables us to give an example to show that sub-cartesian-closedness alone is not sufficient to imply that a functor is sub-logical. Let $\mathcal{D}$ be the monoid of all endomorphisms of the set $\mathbb{N}$ of natural numbers, $\mathcal{C}$ the submonoid of injective endomorphisms, and f the inclusion map. Then f satisfies the condition just given, since any $b : \mathbb{N} \to \mathbb{N}$ can be factored as the composite of an injection a and a split surjection r: for example, we can take $a(n) = 2^{b(n)}(2n+1)$. But the stronger condition of 3.1.2 fails for any b which is not injective.

We shall have more to say about the difference between the conditions of 3.1.2 and of 3.1.3 in the next section (see 3.2.4 and 3.2.14).

Sub-logical and sub-cartesian-closed functors have straightforward stability properties under composition:

Lemma 3.1.4 *Let $F : \mathcal{E} \to \mathcal{F}$ and $G : \mathcal{F} \to \mathcal{G}$ be two cartesian functors between toposes.*

(i) *If both F and G are sub-logical (resp. sub-cartesian-closed), then so is the composite GF.*

(ii) *If GF is sub-logical (resp. sub-cartesian-closed) and G is faithful, then F is sub-logical (resp. sub-cartesian-closed).*

(iii) *If GF is sub-logical (resp. sub-cartesian-closed) and F is logical (resp. cartesian closed) and essentially surjective on objects, then G is sub-logical (resp. sub-cartesian-closed).*

Proof Let A be an object of $\mathcal{E}$. Then the comparison map $\phi_A^{GF} : GF(PA) \to P(GFA)$ is easily seen to be the composite

$$GF(PA) \xrightarrow{G(\phi_A^F)} G(P(FA)) \xrightarrow{\phi_{FA}^G} P(GFA),$$

from which the 'sub-logical' cases of the three statements follow immediately. The 'sub-cartesian-closed' cases follow from the similar factorization of the exponential comparison map $\theta_{A,B}^{GF} : GF(B^A) \to GFB^{GFA}$. □

Next, we note a useful extension of Proposition A4.5.1.

Lemma 3.1.5 *Let j be a local operator on a topos $\mathcal{E}$, and let $L: \mathcal{E} \to \mathbf{sh}_j(\mathcal{E})$ denote the associated sheaf functor. The following conditions are equivalent:*

(i) *j is open.*
(ii) *L is a logical functor.*
(iii) *L is a cartesian closed functor.*
(iv) *L is a sub-logical functor.*
(v) *L is a sub-cartesian-closed functor.*
(vi) *If $m: A' \rightarrowtail A$ is j-dense, so is $m^B: A'^B \rightarrowtail A^B$ for any B.*

Proof The equivalence of (i), (ii), (iii) and (vi) was shown in A4.5.1. The implications (ii) $\Rightarrow$ (iv) and (iii) $\Rightarrow$ (v) are trivial, and (iv) $\Rightarrow$ (v) was shown in 3.1.1 above. So it remains to prove that (v) implies (vi). But since the functor $(-)^B$ preserves pullbacks, it suffices to verify (vi) for the generic j-dense monomorphism $d: 1 \rightarrowtail J$. And $L(J) \cong 1$, so (v) implies that we have a monomorphism $\theta_{B,J}: L(J^B) \rightarrowtail LJ^{LB} \cong 1^{LB} \cong 1$ for any B. So $L(d^B)$ must be an isomorphism; i.e. d^B is j-dense. $\square$

Now we embark on establishing the connection between the foregoing ideas and our previous definition of open map.

Proposition 3.1.6 *Let $F: \mathcal{E} \to \mathcal{F}$ be a cartesian functor between toposes which commutes with universal quantification. Then F is sub-logical.*

Proof Let $A \in \text{ob}\, \mathcal{E}$. Consider the diagram

$$\begin{tikzcd}
F(\Omega^A \times \Omega^A) \arrow[r, "F(\Leftrightarrow^A)"] \arrow[d, "\theta \times \theta"] & F(\Omega^A) \arrow[r, "F(\forall(A))"] \arrow[d, "\theta"] & F\Omega \arrow[dd, "\phi_1"] \\
F\Omega^{FA} \times F\Omega^{FA} \arrow[r, "F(\Leftrightarrow)^{FA}"] \arrow[d, "\phi_1^{FA} \times \phi_1^{FA}"] & F\Omega^{FA} \arrow[d, "\phi_1^{FA}"] & \\
\Omega^{FA} \times \Omega^{FA} \arrow[r, "\Leftrightarrow^{FA}"] & \Omega^{FA} \arrow[r, "\forall(FA)"] & \Omega
\end{tikzcd}$$

in which the right-hand cell commutes by A2.3.11, and the lower left cell commutes by a similar argument using A1.4.13. Since $\Omega^A \times \Omega^A \cong (\Omega \times \Omega)^A$, and similarly for $F\Omega^{FA}$, we can regard the upper left square as an instance of the naturality of θ. So the whole diagram commutes. But the bottom row classifies the diagonal subobject of $\Omega^{FA} \times \Omega^{FA}$, and the top row is similarly the image under F of the classifying map of the diagonal subobject of $\Omega^A \times \Omega^A$; so the

commutativity of the outside of the diagram is tantamount to saying that

$$\begin{tikzcd} F(\Omega^A) \arrow[r, "\phi_A"] \arrow[d, tail, "F(\Delta)"'] & \Omega^{FA} \arrow[d, tail, "\Delta"] \\ F(\Omega^A \times \Omega^A) \arrow[r, "\phi_A \times \phi_A"] & \Omega^{FA} \times \Omega^{FA} \end{tikzcd}$$

is a pullback, i.e. that ϕ_A is monic. □

To obtain a converse to 3.1.6, we need to assume that F has a right adjoint, i.e. that it is the inverse image of a geometric morphism. Then we have

Theorem 3.1.7 *For a geometric morphism $f: \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is open, i.e. f^* is a Heyting functor.*

(ii) *f^* commutes with universal quantification.*

(iii) *f^* is a sub-logical functor.*

(iv) *For any object A of $\mathcal{E}$, the inverse image of the induced geometric morphism $f/A: \mathcal{F}/f^*A \to \mathcal{E}/A$ (cf.* A4.1.3*) is sub-logical.*

(v) *For any object A of $\mathcal{E}$, the inverse image of f/A is sub-cartesian-closed.*

(vi) *For any object A of $\mathcal{E}$ and any subobject B of f^*A, the image of the composite*

$$\mathcal{F}/B \longrightarrow \mathcal{F}/f^*A \xrightarrow{f/A} \mathcal{E}/A$$

is an open subtopos of $\mathcal{E}/A$.

(vii) *The transpose $\overline{\phi_1}: \Omega_{\mathcal{E}} \to f_*\Omega_{\mathcal{F}}$ of ϕ_1 has an internal left adjoint.*

(viii) *For every A, the transpose $\overline{\phi_A}: PA \to f_*(P(f^*A))$ of ϕ_A has an internal left adjoint.*

Proof (i) $\Leftrightarrow$ (ii) since inverse image functors are always coherent, and (ii) $\Rightarrow$ (iii) is 3.1.6 above. For (iii) $\Rightarrow$ (iv), we recall from the proof of A2.3.2 that the power object of an arbitrary object $h: C \to A$ of $\mathcal{E}/A$ is constructed as a subobject of $A^*(PC)$, and it is straightforward to verify that the diagram

$$\begin{tikzcd} (f/A)^*(Ph) \arrow[r, "\phi_h"] \arrow[d, tail] & P(f^*h) \arrow[d, tail] \\ (f/A)^*A^*(PC) \arrow[r, "(f^*A)^*(\phi_C)"] & (f^*A)^*(P(f^*C)) \end{tikzcd}$$

commutes. So ϕ_h is monic if ϕ_C is. (iv) $\Rightarrow$ (v) is immediate from 3.1.1; and (v) $\Rightarrow$ (vi) follows from 3.1.4 and 3.1.5, for if (f/A) has sub-cartesian-closed inverse image then 3.1.4 implies that the same holds for the image of the displayed composite.

(vi) $\Rightarrow$ (viii): Given a morphism $r\colon B \to f_*(P(f^*A))$ in $\mathcal{E}$, corresponding to a subobject $R \rightarrowtail f^*B \times f^*A$ in $\mathcal{F}$, let $\lambda(R)$ be the subobject of $B \times A$ corresponding to the image of the composite geometric morphism $\mathcal{F}/R \to \mathcal{F}/f^*B \times f^*A \to \mathcal{E}/B \times A$, and let $\lambda r\colon B \to PA$ be the name of this subobject. It is easy to see that the construction $R \mapsto \lambda(R)$ is stable under pullback, so that λr is indeed the composite of r with a morphism $\lambda(= \lambda 1)\colon f_*(P(f^*A)) \to PA$. Moreover, given a subobject $S \rightarrowtail B \times A$, named by a morphism $s\colon B \to PA$, it is clear that we have $\lambda(R) \leq S$ in $\mathrm{Sub}(B \times A)$ iff $R \leq f^*S$ in $\mathrm{Sub}(f^*B \times f^*A)$, i.e. that $\lambda r \leq s$ iff $r \leq \overline{\phi_A} s$, yielding the required internal adjunction $(\lambda \dashv \overline{\phi_A})$.

(viii) $\Rightarrow$ (vii) is trivial, so it remains to prove (vii) $\Rightarrow$ (ii). In fact we first show that (vii) implies a strengthened version of (viii): given a left adjoint $\lambda = \lambda_1$ for $\overline{\phi_1}$, we define $\lambda_A\colon f_*(P(f^*A)) \to PA$ to be the composite

$$f_*(\Omega^{f^*A}) \xrightarrow{\theta'} (f_*\Omega)^{f_*f^*A} \xrightarrow{(f_*\Omega)^{\alpha_A}} (f_*\Omega)^A \xrightarrow{\lambda^A} \Omega^A$$

where θ' is the exponential comparison map for f_* and α is the unit of $(f^* \dashv f_*)$. Then it is straightforward to verify that λ_A is an internal left adjoint to $\overline{\phi_A}$, and further that $A \mapsto \lambda_A$ is 'natural' in the sense that, for any $h\colon A \to B$, the diagram

$$\begin{tikzcd} f_*(P(f^*B)) \arrow[r, "\lambda_B"] \arrow[d, "f_*(P(f^*h))"] & PB \arrow[d, "Ph"] \\ f_*(P(f^*A)) \arrow[r, "\lambda_A"] & PA \end{tikzcd}$$

commutes. Taking internal right adjoints of the morphisms in this diagram, we get

$$\begin{tikzcd} PA \arrow[r, "\overline{\phi_A}"] \arrow[d, "\forall h"] & f_*(P(f^*A)) \arrow[d, "f_*(\forall(f^*h))"] \\ PB \arrow[r, "\overline{\phi_B}"] & f_*(P(f^*B)) \end{tikzcd} .$$

But by A2.3.11 this diagram commutes iff f^* preserves universal quantification. $\square$

We digress for a moment to remark that the converse of 3.1.6 holds for direct image functors as well:

Proposition 3.1.8 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f_* commutes with universal quantification.*

(ii) *f_* is sub-logical.*

(iii) *f is an inclusion.*

Proof (iii) $\Rightarrow$ (i): Suppose f is an inclusion, so that we can identify $\mathcal{F}$ with the full subcategory of sheaves for some local operator j in $\mathcal{E}$, and f_* with the inclusion. Given a morphism $g\colon A \to B$ in $\mathbf{sh}_j(\mathcal{E})$ and a subobject $A' \rightarrowtail A$ which is a sheaf, we have to show that $\forall_g(A') \rightarrowtail B$ (as computed in $\mathcal{E}$) is a sheaf. But the subsheaves of a j-sheaf are exactly its j-closed subobjects, by A4.3.8; and since $g^*\colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ commutes with j-closure, it is immediate that its right adjoint $\forall_g$ preserves closed subobjects.

(i) $\Rightarrow$ (ii) again follows from 3.1.6; note that we could also deduce (ii) directly from (iii) using 3.1.1, since the direct images of inclusions are cartesian closed functors by A4.2.9, and the comparison map ϕ_1 in this case is the canonical monomorphism $\Omega_j \rightarrowtail \Omega$.

(ii) $\Rightarrow$ (iii): Suppose f_* is sub-logical. Then it is sub-cartesian-closed by 3.1.1; so, for any two objects A, B of $\mathcal{F}$, any pair of morphisms $1 \rightrightarrows f_*(B^A)$ having equal composites with $\theta_{A,B}\colon f_*(B^A) \rightarrowtail f_*B^{f_*A}$ must be equal. But this says that f_* must be faithful; hence by A4.6.2(*a*) f is localic. And since ϕ_1 is monic, the internal locale in $\mathcal{E}$ which corresponds to f is a sublocale of the terminal locale; so f is an inclusion by 1.5.1(i). □

Given the result of A4.1.12 that any cartesian functor between toposes can be factored as the composite of a direct image functor and an inverse image functor, one might hope to prove the converse of 3.1.6 for arbitrary cartesian functors by combining the results of 3.1.7 and 3.1.8. However, the hypothesis that a functor is sub-logical is not 'stable' under this factorization.

Now we revert to the study of geometric morphisms whose inverse images are Heyting functors. We remark that any geometric morphism whose codomain is a Boolean topos is open: this follows either from the observation after A1.4.10 that a coherent functor defined on a Boolean coherent category is a Heyting functor, or from the result of A4.5.22 that every subtopos of a Boolean topos is open. A more substantial consequence of 3.1.7 is the following:

Corollary 3.1.9

(i) *Any hyperconnected geometric morphism is open.*

(ii) *A geometric morphism is open iff the localic part of its hyperconnected–localic factorization is open.*

Proof (i) By A4.6.6, a morphism $f\colon \mathcal{F} \to \mathcal{E}$ is hyperconnected iff the comparison map $\phi_1'\colon f_*\Omega \to \Omega$ is an isomorphism, and an examination of the proof

of A4.6.6 will show that its inverse is then necessarily $\overline{\phi_1}$. So this is immediate from condition (vii) of 3.1.7.

(ii) follows immediately from (i) and 3.1.4, since the inverse image of a hyperconnected morphism is faithful. □

In connection with 3.1.9(ii), we note also that a geometric morphism is open iff both parts of its surjection–inclusion factorization are open: this follows easily from 3.1.4 and 3.1.5.

It follows from 3.1.9 that the study of open maps of toposes reduces, to a large extent, to the study of open localic maps – equivalently, to the study of open maps of locales, where by 'locales' we understand internal locales in an arbitrary topos. We therefore turn now to the study of such maps, following the approach pioneered by A. Joyal and M. Tierney [560].

We recall the relationship between frames and complete join-semilattices, which we studied in 1.1.10. In particular, we recall that if $f: A \to B$ is a frame homomorphism, then B can be regarded as a commutative monoid in the category of A-modules in **CjSLat**. The connection with open maps is given by

Lemma 3.1.10 *Let $f: X \to Y$ be a map of locales. Then f is open iff $f^*: \mathcal{O}(Y) \to \mathcal{O}(X)$ has a left adjoint $f_!: \mathcal{O}(X) \to \mathcal{O}(Y)$ which is a $\mathcal{O}(Y)$-linear morphism of complete join-semilattices.*

Proof If f^* has a left adjoint $f_!$, the latter is automatically a morphism of complete join-semilattices. The condition that it should be $\mathcal{O}(Y)$-linear is precisely the Frobenius reciprocity condition of 1.5.3(ii). □

As an example of the advantages of this way of viewing open maps, we prove

Proposition 3.1.11 *Let*

$$\begin{tikzcd} P \arrow[r, "k"] \arrow[d, "h"'] & Y \arrow[d, "g"] \\ X \arrow[r, "f"] & Z \end{tikzcd}$$

be a pullback square in **Loc** *with f open. Then*

(i) *k is open;*

(ii) *the square*

$$\begin{tikzcd} \mathcal{O}(X) \arrow[r, "f_!"] \arrow[d, "h^*"'] & \mathcal{O}(Z) \arrow[d, "g^*"] \\ \mathcal{O}(P) \arrow[r, "k_!"] & \mathcal{O}(Y) \end{tikzcd}$$

commutes (we say that the pullback square satisfies the Beck–Chevalley condition*);*

(iii) *if f is also surjective, then k is surjective.*

Proof (i) By 1.1.9, the above pullback in **Loc** corresponds to the pushout

$$\begin{tikzcd} \mathcal{O}(Z) \arrow[r, "f^*"] \arrow[d, "g^*"'] & \mathcal{O}(X) \arrow[d, "h^*"] \\ \mathcal{O}(Y) \arrow[r, "k^*"] & \mathcal{O}(X) \otimes_{\mathcal{O}(Z)} \mathcal{O}(Y) \end{tikzcd}$$

in **Frm**, where $h^*(U) = U \otimes Y$ (recall that we identify Y with the top element of $\mathcal{O}(Y)$) and $k^*(V) = X \otimes V$. By 3.1.10, we know that f^* has an $\mathcal{O}(Z)$-linear left adjoint $f_!$. We define $k_! : \mathcal{O}(X) \otimes_{\mathcal{O}(Z)} \mathcal{O}(Y) \to \mathcal{O}(Y)$ by

$$k_!(U \otimes V) = g^* f_!(U) \cap V;$$

this is well-defined since $(U, V) \mapsto g^* f_!(U) \cap V$ is clearly $\mathcal{O}(Z)$-bilinear. Moreover, $k_!$ is left adjoint to k^* since we have $k_! k^*(V) = g^* f_!(X) \cap V \leq V$ and

$$k^* k_!(U \otimes V) = X \otimes (g^* f_!(U) \cap V) = (X \cap f^* f_!(U)) \otimes V \geq U \otimes V$$

where we used the $\mathcal{O}(Z)$-bilinearity condition at the second step. Finally, $k_!$ is $\mathcal{O}(Y)$-linear, since

$$\begin{aligned} k_!((U \otimes V) \cap k^*(W)) &= k_!((U \otimes V) \cap (X \otimes W)) \\ &= k_!(U \otimes (V \cap W)) \\ &= g^* f_!(U) \cap V \cap W = k_!(U \otimes V) \cap W \end{aligned}$$

for any $U \in \mathcal{O}(X)$ and $V, W \in \mathcal{O}(Y)$. So we have shown that k is open.

(ii) The identity $k_! h^* = g^* f_!$ is clear from the definition of $k_!$.

(iii) If f is a surjection, then $f_!(X) = Z$; so it follows easily that $k_! k^*(V) = g^* f_!(X) \cap V = V$ for all $V \in \mathcal{O}(Y)$, and hence k is a surjection. □

Corollary 3.1.12

(i) *Open surjections are regular epimorphisms in* **Loc**.

(ii) *If an open map is a monomorphism in* **Loc**, *then it is an inclusion (i.e. a regular monomorphism).*

Proof (i) Given an open surjection $f\colon X \to Y$, form the pullback

$$\begin{array}{ccc} P & \xrightarrow{k} & X \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle f} \\ X & \xrightarrow{f} & Y \end{array} \quad ;$$

we must show that f is the coequalizer of h and k in **Loc**, i.e. that

$$\mathcal{O}(Y) \xrightarrow{f^*} \mathcal{O}(X) \underset{k^*}{\overset{h^*}{\rightrightarrows}} \mathcal{O}(P)$$

is an equalizer diagram in **Frm**. But by 3.1.11 k is an open surjection, and the Beck–Chevalley condition holds; so the maps $f_!\colon \mathcal{O}(X) \to \mathcal{O}(Y)$ and $k_!\colon \mathcal{O}(P) \to \mathcal{O}(X)$ give this diagram the structure of a split equalizer in **Set** (in fact in **CjSLat**), and the forgetful functor **Frm** $\to$ **Set** creates equalizers.

(ii) If f is open and a monomorphism, then by (i) and the remark after 3.1.9 the surjective part of its surjection–inclusion factorization is both monic and regular epic, and hence an isomorphism. □

The proof of 3.1.12(i) should be compared with that of A2.2.7. We shall prove a much stronger result – the fact that open surjections are descent morphisms in **Loc** – in 5.1.4.

Remark 3.1.13 If we specialize the Beck–Chevalley condition of 3.1.11 to the case when g is the inclusion of an open sublocale of Z, we recover the Frobenius reciprocity condition of 1.5.3(ii). Thus the assertion 'for every pullback k of f, k^* has a left adjoint $k_!$, and the Beck–Chevalley condition holds for all such pullbacks' is yet another equivalent way of defining what it means for f to be an open map.

We digress briefly to note a couple of results relating openness and fibrewise denseness (cf. 1.2.14), which will be needed in Section C5.3.

Lemma 3.1.14
(i) *Suppose given a pullback square*

$$\begin{array}{ccc} Z' & \xrightarrow{h} & Y' \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle i} \\ Z & \xrightarrow{g} & Y \end{array}$$

in **Loc**$/X$ *for some* X*, where* i *is fibrewise dense over* X *and* g *is an open map. Then* j *is fibrewise dense over* X*.*

(ii) *Suppose given a commutative diagram*

$$\begin{tikzcd} Z \arrow[r, rightarrowtail, "i"] \arrow[dr, "g"'] & Y \arrow[d, "f"] \\ & X \end{tikzcd}$$

where i *is fibrewise dense over* X*. Then* f *is open iff* g *is open. Moreover, if these conditions are satisfied, then the pullback of* i *along an arbitrary locale map* $a\colon X' \to X$ *is fibrewise dense over* X'*.*

Proof (i) The Beck–Chevalley condition of 3.1.11 yields the equality $g^*i_* = j_*h^*\colon \mathcal{O}(Y') \to \mathcal{O}(Z)$, by taking right adjoints. To say that i is fibrewise dense over X means that $i_*i^*f^* = f^*$ (where f is the structure map $Y \to X$); so we immediately deduce $j_*j^*g^*f^* = j_*h^*i^*f^* = g^*i_*i^*f^* = g^*f^*$.

(ii) First suppose g is open. Since i is fibrewise dense, we have $f^* = i_*g^*$; so f^* has a left adjoint $g_!i^*$. To verify the Frobenius reciprocity condition, let $U \in \mathcal{O}(X)$, $V \in \mathcal{O}(Y)$. Then we have

$$\begin{aligned} f_!(f^*(U) \cap V) &= g_!i^*(f^*(U) \cap V) \\ &= g_!(g^*(U) \cap i^*(V)) \\ &= U \cap g_!i^*(V) \\ &= U \cap f_!(V), \end{aligned}$$

as required.

In the case when f is assumed open, we define $g_! = f_!i_*$. Then for $U \in \mathcal{O}(X)$ and $W \in \mathcal{O}(Z)$ we have

$$\begin{aligned} g_!(W) \leq U &\Leftrightarrow f_!i_*(W) \leq U \\ &\Leftrightarrow i_*(W) \leq f^*(U) = i_*g^*(U) \\ &\Leftrightarrow W \leq g^*(U) \end{aligned}$$

using the injectivity of i_* at the final step. The verification of the Frobenius reciprocity condition is similar to the previous case.

For the second assertion of (ii), suppose given $a\colon X \to X'$. We may identify $\mathcal{O}(Y \times_X X')$ with the tensor product $\mathcal{O}(Y) \otimes_{\mathcal{O}(X)} \mathcal{O}(X')$ of $\mathcal{O}(X)$-modules in **CjSLat**, by 1.1.9; in terms of this description, we have to show that if $U \in \mathcal{O}(X')$ and $W \in \mathcal{O}(Y \times_X X')$ are such that $(i^* \otimes 1)(W) \leq Z \otimes U$, then $W \leq Y \otimes U$. But W may be written as a join of tensors $V \otimes U'$, so it suffices to prove the result for

a single such tensor, i.e. that $i^*(V) \otimes U' \leq Z \otimes U$ implies $V \otimes U' \leq Y \otimes U$. Now the projection $Z \times_X X' \to X'$ is open, and by the proof of 3.1.11 the inequality $i^*(V) \otimes U' \leq Z \otimes U$ is equivalent to $a^* g_! i^*(V) \cap U' \leq U$ in $\mathcal{O}(X')$. Since i is fibrewise dense, we have $g_! i^* = f_!$; so this inequality is in turn equivalent to $a^* f_!(V) \cap U' \leq U$, and hence to $V \otimes U' \leq Y \otimes U$ in $\mathcal{O}(Y \times_X X')$. □

Next, we note an important characterization of local homeomorphisms in **Loc**.

Lemma 3.1.15 *A morphism $f: X \to Y$ in* **Loc** *is a local homeomorphism iff both f and the diagonal map $X \to X \times_Y X$ are open. In particular, a locale X is discrete iff both $X \to 1$ and the diagonal $X \to X \times X$ are open maps.*

Proof The second assertion is the case $Y = 1$ of the first, by the remark after 1.6.3. (On the other hand, the first could be deduced by applying the second to internal locales in the topos **Sh**(Y).) It therefore suffices to prove the first.

If f is a local homeomorphism, then it is clearly an open map, since any open $U \rightarrowtail X$ can be covered by opens V for which $f_! V$ is open in Y (and homeomorphic to V), and $f_!$ preserves unions. On the other hand, the pullback $X \times_Y X$ is also a pullback in the category **LH**, by 1.3.2(iv); so the diagonal $X \to X \times_Y X$ is also a local homeomorphism, and hence open.

Conversely, suppose f and the diagonal are both open. Let us call an open $U \in \mathcal{O}(X)$ *small* if $U \otimes U \leq \Delta$ in $\mathcal{O}(X \times_Y X) \cong \mathcal{O}(X) \otimes_{\mathcal{O}(Y)} \mathcal{O}(X)$, where Δ is the element of this frame corresponding to the diagonal. We know that $\mathcal{O}(X) \otimes_{\mathcal{O}(Y)} \mathcal{O}(X)$ is generated as a complete join-semilattice by the elements $U \otimes V$, so we have

$$\Delta = \bigvee \{U \otimes V \mid U, V \in \mathcal{O}(X),\ U \otimes V \leq \Delta\};$$

but $U \otimes V \leq \Delta$ implies $U \otimes V = V \otimes U = (U \cap V) \otimes (U \cap V)$, since the diagonal is the equalizer of the identity on $X \times_Y X$ and the twist map. So

$$\Delta = \bigvee \{U \otimes U \mid U \in \mathcal{O}(X),\ U \otimes U \leq \Delta\}\,.$$

Applying $(\pi_1)_!$, we obtain

$$\begin{aligned} X = (\pi_1)_!(\Delta) &= \bigvee \{(\pi_1)_!(U \otimes U) \mid U \text{ small}\} \\ &\leq \bigvee \{U \in \mathcal{O}(X) \mid U \text{ small}\} \end{aligned}$$

since $U \otimes U \leq U \otimes X = \pi_1^*(U)$ for any U, small or not. Now for any $U \in \mathcal{O}(X)$ the composite $U \rightarrowtail X \to Y$ is open; but the assertion that U is small says precisely that this map is a monomorphism in **Loc**, and hence it is an open inclusion by 3.1.12(ii). So we have shown that X can be covered by opens U for which $U \rightarrowtail X \to Y$ is an open inclusion; but this was how we defined local homeomorphisms in Section C1.3. □

For an internal locale X in a Boolean topos, the unique continuous map $X \to 1$ is always open, by the remark before 3.1.9. In a non-Boolean topos this is not so; however, the Frobenius reciprocity condition is automatic in this case. The reason is that the $\mathcal{O}(1)$-linearity of the left adjoint $X_! : \mathcal{O}(X) \to \mathcal{O}(1)$ of X^* is simply part of what it means to say that $X_!$ is a morphism of internal posets in the topos $\mathcal{S}$ in which X lives (equivalently, that its externalization in the sense of B2.3.3 is an $\mathcal{S}$-indexed functor). Alternatively, we could derive it from 1.2.17(ii), since the terminal locale 1 is regular: note that the existence of a left adjoint for $f^* : \mathcal{O}(Y) \to \mathcal{O}(X)$ says that for each open $U \rightarrowtail X$ there is a smallest open $V \rightarrowtail Y$ through which $U \rightarrowtail X \to Y$ factors, but if every sublocale of Y is an intersection of open sublocales then the latter must coincide with the image of U under f.

It is natural to say that X is an *open locale* if $X \to 1$ is an open map, even though this requires us to take care in distinguishing between open sublocales of a locale X and sublocales of X which are open locales 'in their own right'. (The former implies the latter if X itself is an open locale, but in general there is no implication in either direction.)

We note in passing that a locale X is spatial iff it is a surjective image in $\mathbf{Loc}(\mathcal{S})$ of a discrete locale. Thus the openness of discrete locales, combined with 3.1.4(ii), yields a result which we observed for localic toposes over **Set** in 1.6.6(i):

Corollary 3.1.16 *Every spatial locale is open.* □

Open locales form a useful generalization of the class of spatial locales, which tend to be scarcer in non-Boolean toposes than one might expect. For example, in Section D4.7 we shall encounter the locale of 'formal reals', whose space of points is the Dedekind real number object, and we shall see in D4.7.13 that it is not always spatial; but it is easily seen (from the definition of the formal reals, and from Lemma 3.1.18 below) that it is always open.

If X is a locale, we say an element U of $\mathcal{O}(X)$ is *positive* if every open cover of U is inhabited, i.e. if the sentence

$$(\forall S : P\mathcal{O}(X))((\bigvee S = U) \Rightarrow (\exists V : \mathcal{O}(X))(V \in S))$$

is valid in the internal language of our topos $\mathcal{S}$. (Classically, of course, every element other than the bottom element $\emptyset$ is positive, but constructively this is a stronger condition than 'being nonzero'.) If X is an open locale, it is easy to see that the left adjoint $\lambda_! : \mathcal{O}(X) \to \mathcal{O}(1) = \Omega$ of the unique frame homomorphism $\lambda^* : \Omega \to \mathcal{O}(X)$ sends U to the truth-value of the assertion 'U is positive'. Moreover, we have the following result:

Lemma 3.1.17 *For a locale X, the following are equivalent:*

(i) *X is an open locale.*

(ii) *Every element of $\mathcal{O}(X)$ can be covered by positive elements.*

(iii) *Every covering of an element of $\mathcal{O}(X)$ can be refined to a covering by positive elements.*

(Conditions (ii) and (iii) are of course to be interpreted as sentences in the internal language of the topos in which X lives.)

Proof (i) $\Rightarrow$ (ii): If X is open, then for every $U\colon \mathcal{O}(X)$ we have $U \leq \lambda^*\lambda_!(U)$; but for any $p\colon \Omega$ we have

$$\lambda^*(p) = \bigvee\{V\colon \mathcal{O}(X) \mid (V = X) \wedge p\},$$

since $p = \bigvee\{q : \Omega \mid (q = \top) \wedge p\}$ and λ^* preserves the top element and joins. Thus we obtain $U \leq \bigvee\{V\colon \mathcal{O}(X) \mid (V = X) \wedge \lambda_!(U)\}$, whence $U = \bigvee\{V\colon \mathcal{O}(X) \mid (V = U) \wedge \lambda_!(U)\}$ by the infinite distributive law 1.1.1. But since $\lambda_!(U)$ is the assertion that U is positive, the latter is a covering of U by positive elements.

(ii) $\Rightarrow$ (iii) since, if (ii) holds, we can refine any covering S of U to a covering which is the union of coverings of each of the members of S by positive elements. And (iii) $\Rightarrow$ (ii) follows by applying (iii) to the singleton cover $U = \bigvee\{U\}$.

(ii) $\Rightarrow$ (i): Let us, for the moment, define $\lambda_!(U)$ to be the truth-value of the assertion 'U is positive'; then we have to show that $\lambda_!$ is left adjoint to λ^*. The definition of $\lambda^*(p)$ as a join easily implies that we have

$$(\forall p\colon \Omega)(\lambda_!\lambda^*(p) \Rightarrow p),$$

which yields the counit of the adjunction. For the unit, we note that if U has any covering at all by positive elements then it must be covered by the set $\{V\colon \mathcal{O}(X) \mid (V \leq U) \wedge \lambda_!(V)\}$, and since $\lambda_!$ is order-preserving this is contained in the set $\{V\colon \mathcal{O}(X) \mid (V \leq U) \wedge \lambda_!(U)\}$. And this in turn is a refinement of the set $\{V : \mathcal{O}(X) \mid (V = X) \wedge \lambda_!(U)\}$, whose join is defined to be $\lambda^*\lambda_!(U)$; so the latter must be greater than or equal to U. □

Lemma 3.1.17, sometimes known as the 'Positive Covering Lemma', forms a useful constructive substitute for the classical idea that, in considering open coverings of a set in a topological space, we can restrict our attention to coverings by nonempty open sets. Using it, we may establish an important characterization of open locales in terms of sites:

Lemma 3.1.18

(i) *Let (A, T) be a site whose underlying category is a poset, such that every T-covering family is inhabited. Then the locale corresponding to the frame T-*Idl(A) *of T-ideals of A (cf.* 1.1.16(e)*) is open.*

(ii) *A locale X is open iff $\mathcal{O}(X)$ is isomorphic to T-*Idl(A) *for a site (A, T) as in* (i).

Proof (i) We claim first that, for each $S \subseteq A$, the T-ideal $\langle S\rangle$ generated by S (i.e. the intersection of all T-ideals containing S) has the same support as S

itself (where the support of a set is defined to be the truth-value of the assertion that it is inhabited). To see this, note that for any $p\colon \Omega$ the set $\{a\colon A \mid p\}$ is a T-ideal whose support is contained in p (indeed, equal to p if A is inhabited); but if we take p to be the support of S then it contains S, and hence must contain $\langle S\rangle$. It now follows easily that any inhabited T-ideal is positive as an element of the frame of T-ideals, since the join of a family of T-ideals is the T-ideal generated by their union, and the latter can only be inhabited if the family itself is inhabited. But any T-ideal has a covering by inhabited ones (specifically, $I = \bigvee\{\langle a\rangle \mid a \in I\}$); so by 3.1.17 the locale corresponding to this frame is open.

(ii) One direction is part (i). Conversely, suppose X is open. Let A be the set of positive elements of $\mathcal{O}(X)$, and for each $U \in A$ let $T(U)$ denote the set of subsets of A whose join (in $\mathcal{O}(X)$) is U. Clearly, every member of $T(U)$ is inhabited. But A is a basis for X, by 1.1.17(ii), so we have a bijection between T-ideals in A and arbitrary elements of $\mathcal{O}(X)$, given by $I \mapsto \bigvee I$ and $V \mapsto \{U\colon A \mid U \leq V\}$. (This is really an instance of the Comparison Lemma; cf. 2.2.4(*b*).) □

We remark that the proof of 3.1.18(i) in fact shows that if every T-covering family is inhabited, then the nucleus on the frame LA of lower subsets of A which corresponds to the fixset $T\text{-Idl}(A)$ is strongly dense (i.e. fibrewise dense relative to Ω). Thus the result could alternatively be deduced using 3.1.14(ii). Note also that, in contrast to 2.4.8, 3.1.18(i) does not require the assumption that T satisfies the 'local character' axiom (L). The same is true of its analogue for arbitrary bounded geometric morphisms, although we shall in fact prove the latter under that assumption.

Proposition 3.1.19

(i) *Let $(\mathbb{C}, J)$ be an internal site in a topos $\mathcal{S}$, such that every J-covering family is inhabited. Then the geometric morphism $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J) \to \mathcal{S}$ is open.*

(ii) *A bounded geometric morphism $f\colon \mathcal{E} \to \mathcal{S}$ is open iff there exists a site of definition for $\mathcal{E}$ over $\mathcal{S}$ in which every cover is inhabited.*

Proof (i) We note in passing that, by 2.3.20 and 3.1.9(ii), we may reduce to the case when the underlying category $\mathbb{C}$ is a preorder (so that $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J)$ is localic over $\mathcal{S}$). However, rather than relying on 3.1.18, we shall give a direct proof that the 'constant sheaf' functor $\Delta\colon \mathcal{S} \to \mathbf{Sh}(\mathbb{C}, J)$ is sub-logical.

In fact, using 3.1.1, we shall verify separately that the comparison maps for Ω and for exponentials are monic, rather than working directly with the comparison maps for power objects. First we consider Ω. We note that the condition that all sieves are inhabited ensures that each constant functor $\mathbb{C}^{\mathrm{op}} \to \mathcal{S}$ is separated; hence, to obtain its associated sheaf, we need only apply the '+ construction' of 2.2.6 once, rather than twice. Thus, for an object U of $\mathbb{C}$, an element of $\Delta\Omega(U)$ may be represented by a pair (R, α), where R is a J-covering sieve on U and α assigns to each $f \in R$ a truth-value $\alpha(f) \in \Omega$, subject to the condition that $\alpha(f) = \alpha(fg)$ for every g composable with f (i.e. α is 'constant on connected components of R'). Moreover, two such pairs (R, α) and (S, β) represent the same

element of $\Delta\Omega(U)$ iff there is a covering sieve $T \subseteq R \cap S$ such that the restrictions of α and β to T agree. On the other hand, the subobject classifier of $\mathbf{Sh}(\mathbb{C}, J)$ is the sheaf which sends U to the set $\mathrm{Sv}_J(U)$ of all J-closed sieves on U; and the comparison map $(\phi_1)_U : \Delta\Omega(U) \to \mathrm{Sv}_J(U)$ sends the element represented by (R, α) to the sieve S defined by saying that the truth-value of $(f : V \to U \in S)$ is the truth-value of the assertion that the sieve

$$\{g : W \to V \mid fg \in R \text{ and } \alpha(fg) = \top\}$$

is J-covering. However, we note that if f happens to belong to R then the truth-value of this assertion is precisely $\alpha(f)$ (again, this uses the assumption that every J-covering sieve is inhabited, plus the fact that α is constant on connected components of R); hence if (R, α) and (S, β) are mapped to the same J-closed sieve then α and β must agree on $R \cap S$, and so they define the same element of $\Delta\Omega(U)$. Thus ϕ_1 is monic.

The argument for exponentials is similar. Let A and B be two objects of $\mathcal{S}$; then, as before, elements of $\Delta(B^A)(U)$ are represented by pairs (R, α) where R is a J-covering sieve on U and α is now a function $R \to B^A$ which is constant on connected components. On the other hand, elements of $\Delta B^{\Delta A}(U)$ are natural transformations $\mathbb{C}(-, U) \times A \to \Delta B$ (where we have written A for the constant functor with value A); and the comparison map θ_U sends the element represented by (R, α) to the transformation whose value at a pair $(f : V \to U, a)$ is represented by $(f^*(R), (g \mapsto \alpha(fg)(a)))$. We note that if $f \in R$ then $1_V \in f^*(R)$; in particular $f^*(R)$ is connected, and so the function which forms the second component of the latter representative must be (globally) constant. Hence if θ_U maps (the elements represented by) (R, α) and (S, β) to the same natural transformation, then for any $f \in R \cap S$ and any $a \in A$ we see that $\alpha(f)(a) = \beta(f)(a)$ (this uses the fact that every cover of the domain of f is inhabited), and hence (R, α) and (S, β) represent the same element of $\Delta(B^A)(U)$. So θ is monic, as required.

(ii) Once again, one direction is part (i). For the converse implication, suppose $f : \mathcal{E} \to \mathcal{S}$ is open, and let B be a bound for $\mathcal{E}$ over $\mathcal{S}$. Then from B3.1.6 we know that the $\mathcal{S}$-indexed family of all subobjects of B (indexed by $f_*(PB)$) may be taken as a separating family for $\mathbb{E}$ *qua* $\mathcal{S}$-indexed category. However, for every object A of $\mathcal{E}$, the internal locale $f_*(PA)$ is an open locale in $\mathcal{S}$; this means that (in the internal logic of $\mathcal{S}$) every subobject of B has a covering by positive subobjects (where an object A of $\mathcal{E}$ is positive iff $\mathcal{E}/A \to \mathcal{E} \to \mathcal{S}$ is surjective), and hence that we may alternatively take the family of all positive subobjects of B (indexed by the object I in the pullback

$$\begin{tikzcd}
I \arrow[r, tail] \arrow[d] & f_*(PB) \arrow[d, "\lambda_1"] \\
1 \arrow[r, tail, "\top"] & \Omega_{\mathcal{S}}
\end{tikzcd},$$

where $\lambda_!$ is as usual the left adjoint of the unique frame homomorphism) as a separating family. And if we use the latter to construct a site of definition for $\mathcal{E}$ over $\mathcal{S}$, then it is clear that it will have the property that every cover is inhabited. □

Of course, if $\mathcal{S}$ is Boolean, then every bounded $\mathcal{S}$-topos may be defined by a site with all covers inhabited, as we saw in 2.2.4(e): given an arbitrary site, we simply eliminate those objects which are covered by the empty sieve, and impose the induced coverage on the full subcategory on the objects which remain. But in general this process will not work, unless it happens that the subobject of C_0 consisting of objects with empty covers is complemented, and we also know that every cover is either empty or inhabited.

Before leaving this topic, we note a further refinement of 3.1.19.

Scholium 3.1.20 *A bounded geometric morphism $f: \mathcal{E} \to \mathcal{S}$ is an open surjection iff there is a site of definition $(\mathbb{C}, J)$ for $\mathcal{E}$ over $\mathcal{S}$ such that $\mathbb{C}$ has a terminal object and every J-covering sieve is inhabited.*

Proof One direction follows immediately by combining 3.1.19 with 2.4.8. Conversely, if f is surjective as well as open, then the site of definition for $\mathcal{E}$ constructed in the proof of 3.1.19(ii) will contain the terminal object of $\mathcal{E}$, provided the latter occurs as a subobject of our bound B (and we can ensure the latter by replacing B by $B \amalg 1$, if necessary). □

Corollary 3.1.21 *A geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between bounded $\mathcal{S}$-toposes is an open surjection iff it can be defined by a fibration of sites $(P, T): (\mathbb{D}, K) \to (\mathbb{C}, J)$ in $\mathcal{S}$ for which P preserves covers.*

Proof One direction is immediate from 3.1.20: if we take an internal site for $\mathcal{F}$ over $\mathcal{E}$ satisfying the conditions of 3.1.20 and 'externalize' it as in 2.5.7, we obtain a fibration of sites satisfying the condition given here. Conversely, suppose f is induced by a fibration of sites as above; then by 2.5.10 we already know that f is surjective. We shall show that f^* is sub-logical, using 3.1.1.

We recall that f^* is the composite

$$\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J) \xrightarrow{i_*} [\mathbb{C}^{\mathrm{op}}, \mathcal{S}] \xrightarrow{(-)\circ P} [\mathbb{D}^{\mathrm{op}}, \mathcal{S}] \xrightarrow{j^*} \mathbf{Sh}_{\mathcal{S}}(\mathbb{D}, K)$$

where i_* is the inclusion and j^* the associated sheaf functor. Now the first factor is cartesian closed by A4.2.9, and the second is sub-cartesian-closed since P, being a fibration, satisfies the dual of the condition of 3.1.2. Also, since P preserves covers, it is easy to see that $A \circ P$ is K-separated for any J-sheaf A on $\mathbb{C}$ (cf. the remark after 2.3.3), from which it follows easily that the comparison map $j^*((B \circ P)^{(A \circ P)}) \to j^*(B \circ P)^{j^*(A \circ P)}$ is monic for any A and B. So f^* is sub-cartesian-closed.

It does not seem possible to prove by the same method that the comparison map $f^*(\Omega_{\mathcal{E}}) \to \Omega_{\mathcal{F}}$ is monic, but we can easily do this by direct calculation.

$\Omega_{\mathcal{E}}(P(U,V))$ is the set of J-closed sieves on U, whereas $\Omega_{\mathcal{F}}(U,V)$ is the set of K-closed sieves on (U,V). But if R is any J-closed sieve on U, then the sieve $\{(a,b)\colon (U',V') \to (U,V) \mid a \in R\}$ is easily seen to be K-closed, using the fact that P preserves covers; so the canonical map from the first set of sieves to the second is injective, and applying the associated K-sheaf functor yields the result. □

We remark that 3.1.21 provides a further reason why we did not include the existence and preservation of finite limits in our definition of a fibration of sites (cf. 2.5.9): if we had done so, then the condition in 3.1.21 would imply that P was a morphism of sites, and so it would induce not just a single geometric morphism but an adjoint pair of morphisms as in 2.3.23.

Corollary 3.1.22 *Let*

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

be an inverse sequence of bounded $\mathcal{S}$-toposes (where $\mathcal{S}$ has a natural number object), such that the transition maps f_n are all open surjections. Then the legs $f^n_\infty\colon \mathcal{E}_\infty \to \mathcal{E}_n$ of the limiting cone in $\mathfrak{BTop}/\mathcal{S}$ are open surjections.

Proof We use the technique of 2.5.11 to express $\mathcal{E}_\infty$ as the topos of sheaves on a 'limit site' $(\mathbb{C}_\infty, J_\infty)$ constructed from a sequence of fibrations of sites, each satisfying the condition of 3.1.21. As we noted after 2.5.14, we do not have an explicit description of all the sieves in the coverage J_∞, but the generating sieves which arise from covers in J_n for some n form a coverage (that is, they satisfy the stability axiom (C)), and they are all preserved by P^n_∞, from which it follows that P^n_∞ preserves all covers in J_∞. So the result follows from 3.1.21. □

Next, we consider the stability of open maps and open surjections under pullback in $\mathfrak{Top}$. As with hyperconnected morphisms (cf. 2.4.11(ii)), the stability theorem has two entirely different proofs, depending on which side of the pullback square we assume to be bounded.

One of the two cases follows easily from 3.1.19 and 3.1.20 – or rather, it would follow from these results if we had not assumed in the proof of 3.1.19 that our coverage had local character (cf. 2.4.9). Since we did so, we shall instead use the hyperconnected–localic factorization to reduce to the localic case, for which we can use 3.1.18 (whose proof did not assume local character).

Theorem 3.1.23 *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$, *where* f *is open and bounded. Then* k *is open. If* f *is also surjective, then so is* k.

Proof Since we know that the hyperconnected–localic factorization is stable under pullback (2.4.12), it suffices by 3.1.9 to prove the result when f is not merely bounded but localic. But then we may express $\mathcal{F}$ in the form $\mathbf{Sh}_{\mathcal{E}}(B, T)$, where B is a poset and T is a coverage in which every cover is inhabited; as we observed in 2.4.9, the latter condition is inherited by the pullback coverage $g^{\#}T$, and so k is open by 3.1.18. For the surjective case, we need the extension of 3.1.18 (similar to 3.1.20) where the poset B is assumed to have a top element; but this is straightforward. □

To prove the stability of arbitrary open maps and surjections under bounded pullback, we may of course deal separately with the cases when the bounded morphism is an inclusion and when it has the form $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$. In both cases, the key ingredient is 3.1.7, characterizing open maps as those whose inverse image functors are sub-logical.

Lemma 3.1.24 *Let* $f: \mathcal{F} \to \mathcal{E}$ *be a geometric morphism, and* $\mathbb{C}$ *an internal category in* $\mathcal{E}$. *If* f *is open (resp. surjective), then so is the induced morphism* $[\mathbb{C}, f]: [f^*\mathbb{C}, \mathcal{F}] \to [\mathbb{C}, \mathcal{E}]$.

Proof $[\mathbb{C}, f]^*$ is simply f^* applied to discrete opfibrations over $\mathbb{C}$. It is clear that this is faithful if f^* is, so the 'surjective' case is immediate. The 'open' case is not much harder: if F is a diagram of shape $\mathbb{C}$ in $\mathcal{E}$, corresponding to a discrete opfibration $\mathbb{F} \to \mathbb{C}$, then its power object in $[\mathbb{C}, \mathcal{E}]$ may be computed in $\mathcal{E}$ as a subobject of $PF_1 \times C_0$, equipped with a suitable left $\mathbb{C}$-action, and it is not hard to show that the diagram

$$\begin{tikzcd}
f^*(PF) \arrow[r, "\overline{\phi}_F"] \arrow[d, tail] & P(f^*F) \arrow[d, tail] \\
f^*(PF_1 \times C_0) \arrow[r, "\phi_{F_1} \times 1"] & P(f^*F_1) \times f^*C_0
\end{tikzcd}$$

commutes, where $\overline{\phi}$ is the comparison map for $[\mathbb{C}, f]^*$. Thus if ϕ is monic, so is $\overline{\phi}$. □

To deal with pullbacks of inclusions, we need a result which is of interest in its own right.

Lemma 3.1.25 *Let* $f: \mathcal{F} \to \mathcal{E}$ *be an open geometric morphism; suppose that* j *is a local operator on* $\mathcal{E}$, *and let* k *be the pullback local operator on* $\mathcal{F}$

(cf. A4.5.14(*e*)*). Then there are pullback squares*

$$\begin{tikzcd} f^*(\Omega_j) \arrow[r] \arrow[d, tail] & \Omega_k \arrow[d, tail] \\ f^*(\Omega) \arrow[r, "\phi_1"] & \Omega \end{tikzcd} \quad \textit{and} \quad \begin{tikzcd} f^*(J) \arrow[r] \arrow[d, tail] & K \arrow[d, tail] \\ f^*(\Omega) \arrow[r, "\phi_1"] & \Omega \end{tikzcd}$$

in $\mathcal{F}$ (where the symbols in the top row have their usual meanings).

Proof We recall the construction of k given in A4.5.14(e): it is the smallest local operator such that the composite $f^*(J) \rightarrowtail f^*(\Omega) \to \Omega$ factors through $K \rightarrowtail \Omega$. In particular, we have

$$(\Omega_k \rightarrowtail \Omega) = \forall_{\pi_2}((\pi_1^*(f^*J) \Rightarrow \Theta) \rightarrowtail \Omega \times \Omega)$$

and

$$(K \rightarrowtail \Omega) = \forall_{\pi_1}((\pi_2^*(\Omega_k) \Rightarrow \Theta) \rightarrowtail \Omega \times \Omega)$$

where $\Theta \rightarrowtail \Omega \times \Omega$, as in A4.5.12, is the equalizer of $\Rightarrow$ and π_2. But since f^* is a Heyting functor, it commutes with implication and universal quantification; thus it is easy to see that we have a pullback square

$$\begin{tikzcd} f^*(\Theta_{\mathcal{E}}) \arrow[r] \arrow[d, tail] & \Theta_{\mathcal{F}} \arrow[d, tail] \\ f^*(\Omega \times \Omega) \arrow[r, "\phi_1 \times \phi_1"] & \Omega \times \Omega \end{tikzcd}$$

and hence that the squares in the statement of the lemma are also pullbacks. □

We remark that the conclusion of 3.1.25 actually holds under the weaker assumption that only the comparison map $\phi_1 : f^*(\Omega_{\mathcal{E}}) \to \Omega_{\mathcal{F}}$ is monic. For a proof under this assumption, see [513].

Corollary 3.1.26 *Let f and j be as in* 3.1.25, *and let $f' : \mathbf{sh}_k(\mathcal{F}) \to \mathbf{sh}_j(\mathcal{E})$ be the pullback of f along $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$. Then*

(i) *f' is open;*
(ii) *if f is surjective, so is f'.*

Proof (i) First we note that the comparison map ϕ_1' for the geometric morphism f' is obtained by applying the associated k-sheaf functor

to the morphism $f^*(\Omega_j) \to \Omega_k$ which forms the top edge of the left-hand square in the statement of 3.1.25, since the lemma implies that the square

$$\begin{tikzcd} f^*(1) \arrow[r] \arrow[d, tail, "f^*(\top_j)"] & 1 \arrow[d, tail, "\top_k"] \\ f^*(\Omega_j) \arrow[r] & \Omega_k \end{tikzcd}$$

is a pullback. So we may immediately deduce that ϕ'_1 is monic, since ϕ_1 is. To show the same thing for an arbitrary comparison map ϕ'_A, we recall that $\mathbf{sh}_k(\mathcal{F})$ is an exponential ideal in $\mathcal{F}$ (A4.3.1) and that $B^C \cong B^{LC}$ for any sheaf B, where L is the associated sheaf functor. Using the ideas in the proof of 3.1.1, we may construct ϕ'_A by factoring the composite

$$f^*((\Omega_j)^A) \xrightarrow{\theta} f^*(\Omega_j)^{f^*A} \longrightarrow (\Omega_k)^{f^*A} \cong (\Omega_k)^{f'^*A}$$

through the associated k-sheaf of its domain; but this composite is monic, since both θ and $f^*(\Omega_j) \to \Omega_k$ are, and the associated k-sheaf functor preserves monomorphisms. So we have verified that f'^* is a sub-logical functor.

(ii) Let $m\colon A' \rightarrowtail A$ be a monomorphism in $\mathcal{E}$, with classifying map $\chi\colon A \to \Omega$. Since the right-hand square in the statement of 3.1.25 is a pullback, we see that f^*m is k-dense iff $f^*\chi$ factors through $f^*J \rightarrowtail f^*\Omega$ (equivalently, iff the pullback of $f^*J \rightarrowtail f^*\Omega$ along $f^*\chi$ is an isomorphism). But if f^* is conservative, this happens iff χ factors through $J \rightarrowtail \Omega$, i.e. iff m is j-dense. Since the k-dense monomorphisms are exactly those inverted by the associated sheaf functor $\mathcal{F} \to \mathbf{sh}_k(\mathcal{F})$, it now follows by A1.2.4 that f'^* is conservative. □

Putting together 3.1.24 and 3.1.26, we have established the second version of the pullback-stability theorem:

Theorem 3.1.27 *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$, *where f is open and g is bounded. Then k is open. If f is also surjective, then so is k.* □

Finally (!) in this section, we may establish the connection between openness and the Beck–Chevalley conditions of 2.4.16:

Theorem 3.1.28 *A geometric morphism is stably left weak Beck–Chevalley (in the sense of* 2.4.16*) iff it is open.*

Proof First suppose $f\colon \mathcal{F} \to \mathcal{E}$ is a SLWBC morphism. Then, for any morphism $g\colon B \to A$ in $\mathcal{E}$, the weak Beck–Chevalley condition holds for the square

$$\begin{tikzcd} \mathcal{F}/f^*B \arrow[r, "f^*g"] \arrow[d, "f/B"] & \mathcal{F}/f^*A \arrow[d, "f/A"] \\ \mathcal{E}/B \arrow[r, "g"] & \mathcal{E}/A \end{tikzcd} \quad ;$$

i.e. the canonical natural transformation $(f/A)^*\Pi_g \to \Pi_{f^*g}(f/B)^*$ is monic. Combining this with the natural isomorphism $(f/B)^*g^* \cong (f^*g)^*(f/A)^*$, and recalling that the composite $\Pi_g g^*$ is the exponential functor $(-)^g\colon \mathcal{E}/A \to \mathcal{E}/A$, we deduce that $(f/A)^*$ is sub-cartesian-closed, for any object A of $\mathcal{E}$. But we showed that this is equivalent to openness of f in 3.1.7(v).

Conversely, suppose $f\colon \mathcal{F} \to \mathcal{E}$ is open. Since openness is stable under pullback, it suffices to show that f is a LWBC morphism. But if we consider the pullback of f along an arbitrary bounded morphism $g\colon \mathcal{G} \to \mathcal{E}$, then on forming the hyperconnected–localic factorizations of both f and g we may 'factor' the pullback square into four quarters, as in the proof of 2.4.13; and in three of them the weak Beck–Chevalley condition holds by A4.6.8. Thus we are reduced to the case when both f and g are localic (and f is still open, by 3.1.9). If we now replace them by the corresponding internal locales in $\mathcal{E}$, then 3.1.11(ii) tells us that the (strong) Beck–Chevalley condition holds for subterminal objects of the toposes involved (actually for the left adjoints of the functors we are now considering, but the result we want follows from uniqueness of adjoints). Since our hypotheses are 'stable under slicing' by an object of $\mathcal{E}$, it follows that we have verified condition (iii) of A4.1.17, so by the latter result the weak Beck–Chevalley condition holds for arbitrary objects of the toposes under consideration. □

Suggestions for further reading: Johnstone [513, 523], Joyal & Tierney [560].

C3.2 Proper maps

Having studied open maps of locales and toposes, the reader might reasonably expect us next to turn our attention to closed maps. We begin with closed maps

of locales, which were deliberately omitted from Section C1.5:

Lemma 3.2.1 *For a map $f: X \to Y$ of locales, the following are equivalent:*

(i) *For every closed sublocale C of X, the image of the composite $C \rightarrowtail X \to Y$ is a closed sublocale of Y.*

(ii) *For all $U \in \mathcal{O}(X)$ and $V \in \mathcal{O}(Y)$, we have $f_*(U \cup f^*(V)) = f_*(U) \cup V$, where $f_*: \mathcal{O}(X) \to \mathcal{O}(Y)$ denotes the right adjoint of f^*.*

Proof Since closed sublocales are complementary to open sublocales (1.1.16(*b*)), it is clear that, for any closed sublocale $\complement U$ of X, $\complement f_*(U)$ is the smallest closed sublocale of Y through which the composite $\complement U \rightarrowtail X \to Y$ factors – i.e. it is the closure of the image of this composite. Condition (i) is thus equivalent to saying that, for every such U, the induced locale map $\complement U \to \complement f_*(U)$ is epic. But, if we identify $\mathcal{O}(\complement U)$ with the principal filter $\uparrow(U)$ in $\mathcal{O}(X)$, the frame homomorphism corresponding to this map sends an open $V \geq f_*(U)$ in $\mathcal{O}(Y)$ to $U \cup f^*(V)$, and its right adjoint is simply f_* restricted to elements above U. So condition (ii), for a fixed $U \in \mathcal{O}(X)$, is easily seen to be equivalent to the assertion that the unit of this adjunction is an isomorphism. □

Note that condition (ii) of 3.2.1 is precisely the assertion that the adjunction $(f_* \dashv f^*)$ between $\mathcal{O}(X)^{\mathrm{op}}$ and $\mathcal{O}(Y)^{\mathrm{op}}$ satisfies the Frobenius reciprocity condition of A1.5.8. So, if $\mathcal{O}(X)$ and $\mathcal{O}(Y)$ happen to be coframes as well as frames, it is equivalent to saying that f^* preserves 'co-exponentials'.

Remark 3.2.2 If $\mathcal{O}(Y)$ is Boolean, then every locale map $X \to Y$ is closed: this follows from 3.2.1(i) and the fact that every sublocale of Y is closed (cf. 1.1.20(ii)), but it may also be deduced from 3.2.1(ii), since if V has a complement W in $\mathcal{O}(Y)$ then the identity of 3.2.1(ii) is clearly equivalent to $f_*(f^*(W) \Rightarrow U) = (W \Rightarrow f_*(U))$, and this follows easily from the adjunction $(f^* \dashv f_*)$ and the fact that f^* preserves finite meets.

Lemma 3.2.3 *Closed maps of locales are stable under pullback along local homeomorphisms.*

Proof Suppose given a pullback

$$\begin{tikzcd} X' \arrow[r, "f'"] \arrow[d, "p"] & Y' \arrow[d, "q"] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

where f is closed and q is a local homeomorphism. We must show that the condition of 3.2.1(ii) holds for opens $U' \in \mathcal{O}(X')$, $V' \in \mathcal{O}(Y')$. Consider first the case when q is an open inclusion; then we may regard Y' as an element of

$\mathcal{O}(Y)$, and V' and U' may be identified respectively with an element of $\mathcal{O}(Y)$ satisfying $V' \le Y'$, and an element of $\mathcal{O}(X)$ satisfying $U' \le X' = f^*(Y')$. Under this identification, we have $f'_*(U') = f_*(U') \cap Y'$, and hence

$$\begin{aligned} f'_*(U') \cup V' &= (f_*(U') \cap Y') \cup V' \\ &= (f_*(U') \cup V') \cap Y' \quad \text{since } V' \le Y' \\ &= f_*(U' \cup f^*(V')) \cap Y' \quad \text{since } f \text{ is closed} \\ &= f'_*(U' \cup f'^*(V')) \,. \end{aligned}$$

Now consider the general case: then we have a covering of Y' by open sublocales Y_i such that each composite $Y_i \rightarrowtail Y' \to Y$ is an open inclusion. Using the argument above, we may deduce that $(f'_*(U') \cup V') \cap Y_i = f'_*(U' \cup f'^*(V')) \cap Y_i$ for each i; so the result follows by the infinite distributive law in $\mathcal{O}(Y)$. □

We thus define an arbitrary geometric morphism $f: \mathcal{F} \to \mathcal{E}$ to be closed if, for every object A of $\mathcal{E}$ and every closed subtopos $\mathcal{C}$ of $\mathcal{F}/f^*(A)$, the image of the composite $\mathcal{C} \to \mathcal{F}/f^*(A) \to \mathcal{E}/A$ is a closed subtopos of $\mathcal{E}/A$. By 3.2.3, this agrees with our previous definition for maps between localic toposes.

The next (non-constructive!) lemma should be compared with 3.1.2.

Lemma 3.2.4 *Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories. Then the induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is closed iff f satisfies the following condition: given any $b: V \to fU$ in $\mathcal{D}$, there exist morphisms $a: U' \to U$ in $\mathcal{C}$ and $i: V \to fU'$, $r: fU' \to V$ in $\mathcal{D}$ such that $ri = 1_V$ and $(fa)i = b$.*

Proof We note first that the given condition is stable under pullback along discrete opfibrations: given a discrete opfibration $p: \mathcal{D}' \to \mathcal{D}$, corresponding to a functor $P: \mathcal{D} \to \mathbf{Set}$, lifting the diagram formed by the morphisms a, b, r and i to a similar diagram in $\mathcal{D}'$ is equivalent to choosing an element of $P(V)$. Hence, if the condition is satisfied, it suffices to show that the image under the geometric morphism of any closed subtopos of $[\mathcal{C}, \mathbf{Set}]$ is closed in $[\mathcal{D}, \mathbf{Set}]$. But this is easy; for, as we saw in A4.5.5, the closed subtoposes of $[\mathcal{C}, \mathbf{Set}]$ are those of the form $[\mathcal{S}, \mathbf{Set}]$, where $\mathcal{S}$ is a sieve in $\mathcal{C}$. And by A4.2.12(*b*), the image of the composite geometric morphism $[\mathcal{S}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ may be identified with $[\mathcal{T}, \mathbf{Set}]$, where $\mathcal{T}$ is the full subcategory of $\mathcal{D}$ on all objects which are retracts of ones of the form fU, $U \in \text{ob}\,\mathcal{S}$. But the given condition clearly implies that $\mathcal{T}$ is a sieve if $\mathcal{S}$ is; so the image is a closed subtopos of $[\mathcal{D}, \mathbf{Set}]$.

Conversely, suppose $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is closed. Given a morphism $b: V \to fU$ in $\mathcal{D}$, we first slice over the object $\mathcal{D}(V, -)$: that is, we pull back to the morphism $[(V \downarrow f), \mathbf{Set}] \to [V \backslash \mathcal{D}, \mathbf{Set}]$ induced by f. Now consider the closed subtopos of $[(V \downarrow f), \mathbf{Set}]$ corresponding to the principal sieve $\mathcal{S}$ generated by b (i.e. the sieve of all objects $b': V \to fU'$ which admit a morphism to b). This is clearly nonzero, so its image $[\mathcal{T}, \mathbf{Set}]$ (where $\mathcal{T}$, as before, is the full subcategory of $V \backslash \mathcal{D}$ on those objects which are retracts of some $b': V \to fU'$ in $\mathcal{S}$)

is nonzero, and if it is closed (i.e. if $\mathcal{T}$ is a sieve) then $\mathcal{T}$ must contain the initial object $1_V : V \to V$ of $V\backslash\mathcal{D}$. But this says precisely that the given condition is satisfied for the morphism b. □

Note that the condition of 3.2.4 is dual not to the condition of 3.1.2 for openness of a geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$, but to the weaker condition of 3.1.3 for f^* to be sub-cartesian-closed. The explanation is that the condition of 3.1.3, unlike that of 3.2.4, is not stable under pullback along discrete opfibrations (equivalently, the condition of 3.2.4 is not stable under pullback along discrete fibrations), although it is easy to see directly that it does imply that every open subtopos of $[\mathcal{C}, \mathbf{Set}]$ maps to an open subtopos of $[\mathcal{D}, \mathbf{Set}]$. And if we 'stabilize' it under pullback along discrete opfibrations, we obtain a condition equivalent to that of 3.1.2, by 3.1.7(v).

The preceding paragraph indicates one respect in which the class of closed maps is not as well-behaved as that of open maps: closed maps are not in general stable under pullback. Indeed, this failing already manifests itself in the category of locales (and even in the category of spaces). In **Sp**, the remedy for it is well known: we need to restrict our attention to the closed maps which have compact fibres, otherwise known as *proper maps*. We shall therefore devote the bulk of this section to setting up the theory of proper maps for locales and toposes, which is largely due to J. Vermeulen [1203] in the localic case, and to I. Moerdijk and J. Vermeulen [858] in the topos-theoretic one.

There is a remarkable 'duality' between this theory and the Joyal–Tierney theory of open maps, described in the last section. Where the Joyal–Tierney theory exploits the relationship between the categories **Frm** and **CjSLat**, Vermeulen's theory makes use of the relationship between **Frm** and the category **PFrm** of preframes, which we also studied in 1.1.10. By analogy with 3.1.10, we make the following definition:

Definition 3.2.5 A map of locales $f : X \to Y$ is said to be *proper* if $f^* : \mathcal{O}(Y) \to \mathcal{O}(X)$ has a right adjoint $f_* : \mathcal{O}(X) \to \mathcal{O}(Y)$ in **PFrm** which is $\mathcal{O}(Y)$-linear.

Of course, the right adjoint f_* always exists as a map of ordered sets; so the force of Definition 3.2.5 is that it is a map of preframes (i.e. preserves directed joins) and is $\mathcal{O}(Y)$-linear (i.e. satisfies condition (ii) of 3.2.1). In particular, proper maps are closed.

Proposition 3.2.6 *Proper maps, and proper surjections, are stable under pullback in* **Loc**; *proper surjections are regular epimorphisms in* **Loc**, *and proper*

monomorphisms are inclusions. Moreover, if

$$\begin{tikzcd} P \arrow[r, "k"] \arrow[d, "h"] & Y \arrow[d, "g"] \\ X \arrow[r, "f"] & Z \end{tikzcd}$$

is a pullback square in **Loc** *with f (and therefore k) proper, then the Beck–Chevalley square*

$$\begin{tikzcd} \mathcal{O}(X) \arrow[r, "f_*"] \arrow[d, "h^*"'] & \mathcal{O}(Z) \arrow[d, "g^*"'] \\ \mathcal{O}(P) \arrow[r, "k_*"] & \mathcal{O}(Y) \end{tikzcd}$$

commutes.

Proof This is proved exactly like 3.1.11 and 3.1.12, replacing the tensor product $\mathcal{O}(X) \otimes_{\mathcal{O}(Z)} \mathcal{O}(Y)$ of $\mathcal{O}(Z)$-modules in **CjSLat** with the corresponding tensor product $\mathcal{O}(X) \odot_{\mathcal{O}(Z)} \mathcal{O}(Y)$ of $\mathcal{O}(Z)$-modules in **PFrm**, and replacing binary intersections by binary unions (and left adjoints by right adjoints) throughout the argument. □

Remark 3.2.7 T. Plewe [979] studied a common generalization of open and proper surjections of locales, which he called *triquotient maps* (this term was first used for maps of spaces by E. Michael [807]). A locale map $f: X \to Y$ is called a triquotient map if there exists a function $f_\#: \mathcal{O}(X) \to \mathcal{O}(Y)$ which preserves directed joins and satisfies both the Frobenius conditions

$$f_\#(U \cap f^*(V)) = f_\#(U) \cap V \quad \text{and} \quad f_\#(U \cup f^*(V)) = f_\#(U) \cup V$$

for $U \in \mathcal{O}(X)$, $V \in \mathcal{O}(Y)$. (Note that these two equations force f to be an epimorphism, since the first implies $f_\# f^*(V) \leq V$ and the second implies the reverse inequality.) Clearly, if f is an open surjection then $f_!$ satisfies these conditions (the second Frobenius condition follows from surjectivity and the fact that $f_!$ preserves binary joins), and if f is a proper surjection then f_* does so. (Another class of examples is provided by split epimorphisms in **Loc**: if f is a split epimorphism, then we can take $f_\# = s^*$ for any one-sided inverse s for f.) In general, for a triquotient map f there may be many different choices for the 'triquotient assignment' $f_\#$; but Plewe has shown that if we have a pullback square as in 3.2.6 and f is triquotient, then k is also triquotient, and we may choose its triquotient assignment so that the 'Beck–Chevalley' square commutes.

If X is a locale in a Boolean topos, then the unique locale map $\gamma: X \to 1$ is proper iff γ_* preserves directed joins (i.e. is a morphism of preframes), by Remark 3.2.2; but this is exactly how we defined compactness for locales in 1.5.5. Constructively, the Frobenius reciprocity condition in the definition of propriety (unlike that for openness) is a nontrivial restriction even on locale maps with codomain 1; so we include it in the definition of compactness – that is, we *define* X to be compact if $X \to 1$ is proper.

The justification for studying proper maps of locales is the following result, in which the equivalence of (i) and (ii) parallels a well-known result for spaces.

Theorem 3.2.8 *Let $f: X \to Y$ be a map of locales. The following are equivalent:*

(i) *f is proper.*

(ii) *f is stably closed, i.e. its pullback along any morphism with codomain Y is closed.*

(iii) *The internal locale in* $\mathbf{Sh}(Y)$ *corresponding to f under the equivalence of 1.6.3 is compact.*

Proof (i) $\Rightarrow$ (ii) is immediate from 3.2.6 and the fact, already noted, that proper maps are closed.

(ii) $\Rightarrow$ (iii): Condition (ii) is 'stable' under passage across the equivalence of 1.6.3, since the notion of closed sublocale is stable in this sense. Thus it is sufficient to prove the implication (provided we do so constructively) in the case $Y = 1$: that is, given a locale X such that the projection $Z \times X \to Z$ is closed for every Z, we have to show that X is compact. The particular case $Z = 1$ takes care of the Frobenius reciprocity condition, so we have only to verify that the top element of $\mathcal{O}(X)$ is inaccessible by directed joins. Let $\mathcal{D}$ be a directed subset of $\mathcal{O}(X)$ with join X; we construct a 'test locale' Z as follows. As at the end of Section C1.2, we write $d: X_d \to X$ for the dissolution of X; then $d^*(U)$ is closed as well as open in X_d for each $U \in \mathcal{O}(X)$ (in particular, for each $U \in \mathcal{D}$), and we write $e(U)$ for its complement. Let $\mathcal{F}$ be the filter in $\mathcal{O}(X_d)$ generated by $\{e(U) \mid U \in \mathcal{D}\}$, and let $\mathcal{O}(Z)$ be the subframe

$$\{(p, V) \mid p \Rightarrow (V \in \mathcal{F})\}$$

of $\Omega \times \mathcal{O}(X_d)$. (Thus $\mathcal{O}(Z)$ is obtained by applying the frame-theoretic version of the glueing construction (A2.1.12) to the classifying map $\chi: \mathcal{O}(X_d) \to \Omega$ of $\mathcal{F}$: note that χ preserves finite meets precisely because $\mathcal{F}$ is a filter.)

Now consider the element

$$W = \bigvee \{(\top, e(U)) \otimes U \mid U \in \mathcal{D}\}$$

of $\mathcal{O}(Z \times X) \cong \mathcal{O}(Z) \otimes \mathcal{O}(X)$. We have

$$\begin{aligned} W \vee ((\bot, X_d) \otimes X) &= W \vee \bigvee\{(\bot, X_d) \otimes U \mid U \in \mathcal{D}\} \\ &= \bigvee\{((\top, e(U)) \vee (\bot, X_d)) \otimes U \mid U \in \mathcal{D}\} \\ &= \bigvee\{(\top, X_d) \otimes U \mid U \in \mathcal{D}\} = Z \otimes X, \end{aligned}$$

where we have twice used the fact that $\mathcal{D}$ covers X. Since the projection $\pi_1 \colon Z \times X \to Z$ is closed, it follows from 3.2.1(ii) that we have $\pi_{1*}(W) \vee (\bot, X_d) = Z$ in $\mathcal{O}(Z)$. So $\pi_{1*}(W)$ must have the form $(\top, V)$ for some $V \in \mathcal{F}$, and hence $\pi_{1*}(W) \geq (\top, e(U))$ for some $U \in \mathcal{D}$, or equivalently $(\top, e(U)) \otimes X \leq W$. Applying the frame homomorphism $\mathcal{O}(Z) \otimes \mathcal{O}(X) \to \mathcal{O}(X_d)$ which sends $(p, V) \otimes U$ to $V \wedge d^*(U)$, we obtain

$$e(U) \leq \bigvee\{e(U') \wedge d^*(U') \mid U' \in \mathcal{D}\} = \emptyset\,.$$

But this implies that $d^*(U)$ is the whole of X_d, and hence (since d is a surjection) that U is the whole of X.

(iii) $\Rightarrow$ (i): Suppose f is such that $\mathcal{I}(f)$ is a compact internal locale in $\mathbf{Sh}(Y)$. Since, by 1.6.2, the inverse of $\mathcal{I}$ is simply the global sections functor applied to internal frames in $\mathbf{Sh}(Y)$, it is easy to see that f inherits the Frobenius reciprocity condition of 3.2.1(ii) from the unique morphism $\mathcal{I}(f) \to 1$ in $\mathbf{Loc}(\mathbf{Sh}(Y))$. So we have only to prove that f_* preserves directed joins.

Let $\mathcal{D}$ be a directed subset of $\mathcal{O}(X)$; it will be convenient to suppose that $\mathcal{D}$ is also downwards closed, and hence an ideal. Recalling that $\mathcal{O}(\mathcal{I}(f))$ is the sheaf on Y whose value at $V \in \mathcal{O}(Y)$ is $\mathcal{O}(f^*V)$, we define an ideal $\widetilde{\mathcal{D}}$ of this internal lattice in $\mathbf{Sh}(Y)$ by

$$\widetilde{\mathcal{D}}(V) = \{U \in \mathcal{O}(f^*V) \mid \ \{V' \leq V \mid U \cap f^*V' \in \mathcal{D}\} \text{ covers } V\}\,.$$

It is clear that this a sheaf; moreover, computation of the truth-value of the assertion '$\widetilde{\mathcal{D}}$ covers $\mathcal{I}(f)$' in $\mathbf{Sh}(Y)$ yields the element $V_0 = f_*(\bigvee \mathcal{D})$ of $\mathcal{O}(Y)$. So the compactness of $\mathcal{I}(f)$ implies that $f^*(V_0) \in \widetilde{\mathcal{D}}(V_0)$, i.e. that V_0 can be covered by opens V' whose inverse images under f are in $\mathcal{D}$. But for each such V' we have $V' \leq f_*f^*(V') \leq V_0$; hence V_0 is also the join of $\{f_*(U) \mid U \in \mathcal{D}\}$, i.e. f_* preserves the join of $\mathcal{D}$. $\square$

Remark 3.2.9 An inclusion in **Loc** is proper iff it is closed in the sense of 3.2.1, iff it is the inclusion of a closed sublocale as defined in 1.2.6(*b*). The latter equivalence is immediate from 3.2.1(i); for the former, we note that if $f \colon \complement U \to X$ is a closed inclusion, then f_* maps $\mathcal{O}(\complement U)$ onto the principal filter $\uparrow(U) \subseteq \mathcal{O}(X)$, and so preserves all nonempty joins. (Alternatively, we could use 1.2.10 and 3.2.8(ii).) Thus the 'proper' analogue of the class of discrete locales (3.1.15) is the class of compact Hausdorff locales (cf. 1.2.17).

We should also remark that, just as there is a (largely obsolete) tradition in general topology whereby 'compact' means 'compact Hausdorff', and the term 'quasicompact' is used for spaces which satisfy the open-covering definition of compactness but are not necessarily Hausdorff, so there is a tradition – which has been followed by some writers on locales and/or toposes – whereby 'proper map' means one which is proper in our sense and also has closed (i.e. proper) diagonal; authors who adhere to this tradition use the word 'perfect' to describe the maps which we call proper. However, 'perfect' is also used by some authors to denote a *stronger* condition than properness, so it seems safest to avoid the word altogether. Thus we shall say that a locale morphism $f: Y \to X$ is *proper and separated* if it corresponds to a compact Hausdorff internal locale in $\mathbf{Sh}(X)$, i.e. if it is proper and the diagonal $Y \rightarrowtail Y \times_X Y$ is closed.

Compactness has its expected properties in Hausdorff locales:

Lemma 3.2.10

(i) *Any continuous map from a compact locale to a Hausdorff locale is proper. In particular, any compact sublocale of a Hausdorff locale is closed.*

(ii) *A compact locale X is Hausdorff iff it is regular; that is, each $U \in \mathcal{O}(X)$ can be expressed as a union of opens whose closures are contained in U.*

Proof (i) Suppose $f: Y \to X$ is a locale map such that Y is compact and X is Hausdorff. Then $(1, f): Y \rightarrowtail Y \times X$ is a closed inclusion, because it is the pullback of the diagonal $\Delta: X \rightarrowtail X \times X$ along $f \times 1$. But the projection $\pi_2: Y \times X \to X$ is proper (because it is a pullback of $Y \to 1$) and hence $f = \pi_2(1, f)$ is proper. The second assertion is the special case of the first when f is an inclusion, by 3.2.9.

(ii) One direction was proved (without the assumption of compactness) in 1.2.17. For the converse, suppose X is compact and Hausdorff, and let U be open in X. The complement $\complement U$ is compact (as is any closed sublocale of a compact locale), and so the projection $\pi_2: \complement U \times U \to U$ is proper by 3.2.6; equivalently, by 3.2.8, it corresponds to a compact internal locale $\mathcal{I}(\pi_2)$ in $\mathbf{Sh}(U)$. But $\complement U \times U$ is disjoint from the diagonal sublocale of $X \times X$, so it is contained in the open complement of the latter, i.e. the union of all open rectangles $P \times Q \rightarrowtail X \times X$ such that $P \cap Q = 0$. Now $\mathcal{O}(\mathcal{I}(\pi_2))$ is the internal frame $(V \mapsto \mathcal{O}(\complement U \times V))$. The sub-presheaf F of this given by

$$V \mapsto \{Z \times V \mid Z = P \cap \complement U \text{ for some } P \in \mathcal{O}(X) \text{ with } P \cap V = 0\}$$

is not a sheaf; but we may form its closure $\overline{F}$ in the sheaf $\mathcal{O}(\mathcal{I}(\pi_2))$ already mentioned, that is the sheaf whose sections over $V \in \mathcal{O}(U)$ are those open sublocales W of $\complement U \times V$ such that V has an open covering by sublocales V_i for which $W \cap (\complement U \times V_i)$ is of the form $(P_i \cap \complement U) \times V_i$ for some $P_i \in \mathcal{O}(X)$ disjoint from V_i. It is easily seen that $\overline{F}$ is directed in the ordering inherited from $\mathcal{O}(\mathcal{I}(\pi_2))$ (indeed, it is closed under finite joins in the latter), and that its join is the top element $\complement U \times U$. So by compactness of $\mathcal{I}(\pi_2)$ we must have $\complement U \times U \in \overline{F}(U)$; but

this says that U can be covered by opens V for which $\complement U \leq \neg V$ – equivalently, for which the closure of V is contained in U. So X is regular. $\square$

One further result on proper maps of locales which we shall need is the following:

Proposition 3.2.11 *Suppose given a diagram of locales $(X_i \mid i \in \mathrm{ob}\,\mathcal{I})$ indexed by a cofiltered category $\mathcal{I}$, such that the transition maps $f_\alpha \colon X_j \to X_i$ induced by morphisms $\alpha \colon j \to i$ in $\mathcal{I}$ are all proper. Then the legs $g_j \colon X_\infty = \varprojlim_{\mathcal{I}} X_i \to X_j$ of the limit cone in* **Loc** *are all proper.*

Proof First suppose that the transition maps are surjective as well as proper. By 1.1.8 and 1.1.10, we know that the filtered colimit $\mathcal{O}(X_\infty)$ in **Frm** is also a colimit in **PFrm**; now, for each j, the right adjoints $f_{\alpha *} \colon \mathcal{O}(X_i) \to \mathcal{O}(X_j)$ form a cone in **PFrm** under the $\mathcal{I}/j$-indexed diagram formed by the $\mathcal{O}(X_i)$ with morphisms $\alpha \colon i \to j$, since $(f_{\alpha\beta})_* f_\beta^* \cong f_{\alpha *} f_{\beta *} f_\beta^* \cong f_{\alpha *}$ by surjectivity. Since $\mathcal{I}$ is cofiltered, the forgetul functor $\mathcal{I}/j \to \mathcal{I}$ is initial, and so this diagram has the same colimit as the original one; so we get a preframe homomorphism $\mathcal{O}(X_\infty) \to \mathcal{O}(X_j)$ for each j, which is easily verified to be the right adjoint g_{j*} of g_j^*, and to be $\mathcal{O}(X_j)$-linear (since all the $(f_\alpha)_*$ are $\mathcal{O}(X_j)$-linear). So g_j is proper (and surjective – though we already knew the latter fact, from 1.1.12).

Now consider the general case. We may replace each X_i by the closed sublocale C_i which is the intersection, over all $\alpha \colon j \to i$, of the image of $f_\alpha \colon X_j \to X_i$; equivalently,

$$C_i = \complement\left(\bigcup \{ f_{\alpha *}(\emptyset) \mid \alpha \colon j \to i \}\right) .$$

Clearly, each $g_i \colon X_\infty \to X_i$ must factor through $C_i \rightarrowtail X_i$, so this change does not affect the limit. Also, since the union appearing in the definition of the open complement of C_i is directed, it is preserved by the right adjoints of the transition maps; hence, since they are closed maps, their restrictions to maps $C_j \to C_i$ are surjective as well as proper. So, by the particular case already considered, each $X_\infty \to C_i$ is proper; composing with the closed inclusion $C_i \rightarrowtail X_i$ yields the result. $\square$

In fact a stronger result than 3.2.11 is true: if the vertices X_i of our cofiltered diagram are compact locales, then the limit X_∞ is compact, even if the transition maps fail to be proper. However, the proof of this stronger result requires a more complicated argument, and we shall not give it here – we refer the interested reader to [435].

In seeking to generalize the notion of proper map from locales to toposes, we have to decide whether our definition of 'proper map of toposes' should include all hyperconnected maps. We have seen in the last section that the natural definition of 'open map' for toposes does include all hyperconnected maps, and hence that a map is open iff the localic part of its hyperconnected–localic factorization

is open. Thus, if we wish to preserve the 'duality' between openness and properness which we have observed at the localic level, we should expect the same thing to be true for proper maps: indeed, we may reasonably *define* a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ to be proper if the localic part of its hyperconnected–localic factorization is proper, i.e. if $f_*(\Omega_{\mathcal{F}})$ is a compact internal frame in $\mathcal{E}$.

What does it mean to say that an internal frame (in particular, one of the form $f_*(\Omega_{\mathcal{F}})$) is compact in a general topos $\mathcal{E}$? For the case $\mathcal{E} = \mathbf{Set}$, we saw how to answer this in topos-theoretic terms in 1.5.5: it means that the direct image functor $\mathbf{Sh}(X) \to \mathbf{Set}$ (where X is the locale corresponding to the frame in question) preserves directed colimits of subterminal objects. Another way of saying the same thing is that, if $\mathcal{I}$ is any directed poset, then the canonical natural transformation in the diagram

$$\begin{tikzcd}
{[\mathcal{I}, \mathbf{Sh}(X)]} \arrow[r] \arrow[d, "\varinjlim_{\mathcal{I}}"] \arrow[dr, phantom, "\downarrow"] & {[\mathcal{I}, \mathbf{Set}]} \arrow[d, "\varinjlim_{\mathcal{I}}"] \\
\mathbf{Sh}(X) \arrow[r] & \mathbf{Set}
\end{tikzcd}$$

is an isomorphism when evaluated at a subterminal object of $[\mathcal{I}, \mathbf{Sh}(X)]$. But this is easily recognizable (modulo 'stabilization under slicing') as a weak Beck–Chevalley condition, given that the $\varinjlim_{\mathcal{I}}$ are inverse image functors.

Our formal definition is thus as follows:

Definition 3.2.12 (*a*) We say that a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ is *proper*, or that $\mathcal{F}$ is *compact* as an $\mathcal{E}$-topos, if, for any object A of $\mathcal{E}$ and any directed internal poset $\mathbb{I}$ in $\mathcal{E}/A$, the Beck–Chevalley natural transformation for the square

$$\begin{tikzcd}
\mathcal{F}/f^*A \arrow[r, "f/A"] \arrow[d, "\infty_{\mathbb{I}}"] & \mathcal{E}/A \arrow[d, "\infty_{\mathbb{I}}"] \\
{[\mathbb{I}, \mathcal{F}/f^*A]} \arrow[r, "{[\mathbb{I}, f/A]}"] & {[\mathbb{I}, \mathcal{E}/A]}
\end{tikzcd}$$

(where $\infty_{\mathbb{I}}$ denotes the geometric morphism whose inverse image is $\varinjlim_{\mathbb{I}}$; cf. B2.6.9) is an isomorphism when evaluated at subterminal objects of $[\mathbb{I}, \mathcal{F}/f^*A]$.

(*b*) We say that a bounded geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ is *separated*, or that $\mathcal{F}$ is *Hausdorff* as an $\mathcal{E}$-topos, if the diagonal $\Delta_f\colon \mathcal{F} \to \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ is proper.

In view of B2.6.13, we could replace directed posets by arbitrary filtered categories in this definition; but it seems more natural to use posets here since we are dealing with subterminal objects. We could also eliminate the variable A

from the definition by working with weakly filtered categories (or weakly directed posets), since we saw in B2.6.6 that an internal category $\mathbb{C}$ is weakly filtered iff it is filtered as an internal category in $\mathcal{E}/\pi_0\mathbb{C}$. Note also that the square in the definition is a pullback, since its bottom edge is the pullback of f/A along the canonical morphism $[\mathbb{I}, \mathcal{E}/A] \to \mathcal{E}/A$; and since f/A is itself a pullback of f, the following result is immediate from A4.1.17.

Lemma 3.2.13 *Any stable right weak Beck–Chevalley morphism (as defined in 2.4.16) is proper.* □

In particular, any hyperconnected morphism is proper, by 2.4.17. Once we have proved the stability of proper maps under pullback, we shall obtain a converse to 3.2.13.

Lemma 3.2.14 *Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories. The induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is proper iff*

(i) *For each object V of $\mathcal{D}$, the comma category $(V \downarrow f)$ has a finite solution-set (that is, there exists a finite family of morphisms $b_j: V \to fU_j$ $(1 \leq j \leq n)$ such that every $b: V \to fU$ factors as $(fa)b_j$ for some j and some $a: U_j \to U$ in $\mathcal{C}$); and*

(ii) *For any $b: V \to fU$, there exist $a: U' \to U$, $i: V \to fU'$ and $r: fU' \to V$ such that $ri = 1_V$ and $br = fa$.*

Note that condition (ii) above is stronger than the condition of 3.2.4: it is dual to the condition of 3.1.2, rather than that of 3.1.3. This is not surprising, given the fact that proper maps are stable under pullback (cf. the remarks after 3.2.4). We note also that the two conditions could be combined into one, saying that for any V there exists a finite set of pairs $(i_j: V \to fU_j, r_j: fU_j \to V)$ with $r_j i_j = 1_V$, such that for any $b: V \to fU$ there exists $a: U_j \to U$ for some j such that $br_j = fa$. However, it is usually more convenient to consider the two conditions separately.

Proof First suppose the conditions are satisfied. We note that both of them are stable under pullback along discrete opfibrations (if $p: \mathcal{D}' \to \mathcal{D}$ is a discrete opfibration, $f': \mathcal{C}' \to \mathcal{D}'$ the pullback of f along p and V' an object of $\mathcal{D}'$, then the comma category $(V' \downarrow f')$ is isomorphic to $(pV' \downarrow f)$). So it suffices to prove that for any directed internal poset Q in $[\mathcal{D}, \mathbf{Set}]$ and any f^*Q-indexed family of subterminal objects of $[\mathcal{C}, \mathbf{Set}]$, the functor f_* preserves the join of this family.

Now a directed internal poset in $[\mathcal{D}, \mathbf{Set}]$ is just a functor $Q: \mathcal{D} \to \mathbf{Poset}$ whose values are directed posets; equivalently, we may regard it (via the Grothendieck construction) as a split opfibration $q: \mathcal{Q} \to \mathcal{D}$ whose fibres are directed posets. Similarly, we shall regard $P = f^*(Q): \mathcal{C} \to \mathbf{Poset}$ as a split opfibration $p: \mathcal{P} \to \mathcal{C}$. Now a subterminal object R in $[P, [\mathcal{C}, \mathbf{Set}]] \simeq [\mathcal{P}, \mathbf{Set}]$ is just a cosieve $\mathcal{R}$ in $\mathcal{P}$, and its join in $[\mathcal{C}, \mathbf{Set}]$ is the cosieve in $\mathcal{C}$ which is the image of the composite fibration $\mathcal{R} \to \mathcal{P} \to \mathcal{C}$. Applying f_* to this, we obtain the cosieve

of those $V \in \text{ob}\ \mathcal{D}$ such that, for every $b\colon V \to fU$, U lies in the image of $\mathcal{R}$; equivalently, there exists an index π in $P(U) = Q(fU)$ which belongs to the subset $R(U)$. However, if we first apply the direct image functor $[\mathcal{P}, \mathbf{Set}] \to [\mathcal{Q}, \mathbf{Set}]$ and then compute the join in $[\mathcal{D}, \mathbf{Set}]$, we obtain the cosieve of those V for which there exists an index γ in $Q(V)$ such that, for every $b\colon V \to fU$, we have $Qb(\gamma) \in R(U)$. It is clear that the first cosieve contains the second; we have to show the reverse inclusion.

To do this, suppose V belongs to the first cosieve. We first choose a finite solution set $(b_j : V \to U_j \mid 1 \leq j \leq n)$ as in condition (i), and then for each j we choose a_j, i_j and r_j as in condition (ii). By assumption, we have elements $\pi_j \in P(U'_j) = Q(fU'_j)$ (where U'_j is the domain of a_j) which belong to $R(U'_j)$; let $\gamma_j = Qr_j(\pi_j) \in Q(V)$. Since $Q(V)$ is directed, we may choose a single index γ in it which is an upper bound for the γ_j; then for each j we have $Qb_j(\gamma) \geq Qb_j(\gamma_j) = Pa_j(\pi_j) \in P(U_j)$, and hence $Qb_j(\gamma) \in R(U_j)$ since $\mathcal{R}$ is a cosieve in $\mathcal{P}$. But since the b_j form a solution set for $(V \downarrow f)$, this is sufficient.

To show the necessity of condition (i), let V be an object of $\mathcal{D}$. Pulling back along the discrete opfibration corresponding to $\mathcal{D}(V, -)\colon \mathcal{D} \to \mathbf{Set}$, we obtain the functor $\overline{f}\colon (V \downarrow f) \to V \backslash \mathcal{D}$ induced by f; so the latter preserves directed joins of subterminal objects. Now let P be the (external) poset of finitely-generated cosieves in $(V \downarrow f)$, ordered by inclusion; we may regard it as a constant directed internal poset in both $[V \backslash \mathcal{D}, \mathbf{Set}]$ and $[(V \downarrow f), \mathbf{Set}]$, and in the latter we have a P-indexed family of subterminal objects, namely the elements of P themselves. If we compute the join of this family in $[\mathcal{C}, \mathbf{Set}]$, we obtain the terminal object, since each object of $\mathcal{C}$ belongs to a finitely-generated cosieve; and f_* preserves the terminal object. So the initial object 1_V of $V \backslash \mathcal{D}$ belongs to the join of the cosieves $f_* S$, $S \in P$, and hence belongs to some particular $f_* S$. But this says precisely that S is the whole of $(V \downarrow f)$; i.e. $(V \downarrow f)$ is finitely-generated as a cosieve in itself.

We shall not give a detailed proof of the necessity of (ii), since it follows from two general results about proper maps to be proved later. We shall see in 3.2.17(ii) that proper maps are closed; hence if $f\colon \mathcal{C} \to \mathcal{D}$ induces a proper map between functor categories then it must satisfy the weak version of (ii) considered in 3.2.4. But we shall also see that proper maps are stable under arbitrary pullback (3.2.21); and we observed after 3.2.4 that if we 'stabilize' the weak condition under pullback (specifically, along discrete fibrations), we obtain the stronger condition (ii) above. □

Although the last part of the above proof is non-constructive, the proof of sufficiency of the conditions is not; so they can be used to give a criterion for a functor between internal categories in an arbitrary topos to induce a proper map. (In this context, the notion of finiteness appropriate to condition (i) is, as usual, the Kuratowskian one studied in Section D5.4. In particular, it is not hard to see that a local homeomorphism $\mathcal{E}/A \to \mathcal{E}$ is proper iff A is a K-finite object of $\mathcal{E}$.)

Example 3.2.15 Nontrivial examples of functors satisfying the conditions of 3.2.14 are not all that common, but the following one will be of use in 5.2.9 below. Let $\mathcal{D}$ be the category of finite sets and surjections between them, and let $\mathcal{C}$ be the poset whose elements are finite composable strings $(b_1, b_2, \ldots, b_n)$ of *non-invertible* morphisms of $\mathcal{D}$ (that is, strings such that dom b_j = cod b_{j+1} for all j), ordered by setting $(b_1, b_2, \ldots, b_n) \leq (b'_1, b'_2, \ldots, b'_m)$ iff the latter is an initial segment of the former, i.e. $m \leq n$ and $b'_j = b_j$ for all $j \leq m$. (It is convenient to include strings of length 0, that is objects of $\mathcal{D}$, as maximal elements of $\mathcal{C}$, with $(b_1, b_2, \ldots, b_n) \leq V$ iff V = cod b_1.) We have a functor $f: \mathcal{C} \to \mathcal{D}$ which sends $(b_1, b_2, \ldots, b_n)$ to dom b_n, and an inequality $U' \leq U$ to the composite of the morphisms belonging to the string U' but not to U; we claim that this functor satisfies the conditions of 3.2.14.

The first condition is trivial, because the comma category $(V \downarrow f)$ is in fact finite for each object V of $\mathcal{D}$ (this is where we use the non-invertibility condition in the definition of $\mathcal{C}$). For the second, we must distinguish two cases: if $b: V \to fU$ is invertible, then we set $a = 1_U$, $i = b$ and $r = b^{-1}$, and if not we let $U' \leq U$ be the string obtained by adjoining b to the end of U, and set $i = r = 1_V$.

We now embark on the study of proper morphisms in general.

Lemma 3.2.16 *Suppose given a commutative triangle*

$$\begin{tikzcd} \mathcal{G} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

of geometric morphisms.

(i) *If f and g are proper, then h is proper.*

(ii) *If h is proper and g is a surjection, then f is proper.*

(iii) *If h is proper and f is an inclusion, then g is proper.*

Proof (i) If $\mathbb{I}$ is a directed internal poset in $\mathcal{E}/A$, then $f^*\mathbb{I}$ is a directed internal poset in $\mathcal{F}/f^*A$. So we may patch together the weak Beck–Chevalley squares for f and g, to obtain that for h.

(ii) As before, suppose $\mathbb{I}$ is a directed internal poset in $\mathcal{E}/A$. The morphism $[h^*\mathbb{I}, \mathcal{G}/h^*A] \to [f^*\mathbb{I}, \mathcal{F}/f^*A]$ induced by g is surjective, so any subterminal object V of its codomain satisfies $V \cong g_*g^*V$ (cf. A4.2.6(v)). But this is exactly what we need to deduce the weak Beck–Chevalley condition for f from that for h.

(iii) Here we cannot use the same style of argument, since we do not know that every directed internal poset in $\mathcal{F}$ is the image under f^* of one in $\mathcal{E}$. However,

if we write X and Y for the internal locales in $\mathcal{E}$ corresponding to the frames $h_*(\Omega_{\mathcal{G}})$ and $f_*(\Omega_{\mathcal{F}})$, then we have a pullback diagram in $\mathbf{Loc}(\mathcal{E})$, since $Y \to 1$ is

$$\begin{tikzcd} X \arrow[r, "1"] \arrow[d] & X \arrow[d] \\ Y \arrow[r] & 1 \end{tikzcd}$$

monic and $X \to 1$ factors through it. So $X \to Y$ is a proper map of internal locales, by 3.2.6; but this says that $g_*(\Omega_{\mathcal{G}})$ is a compact internal locale in $\mathcal{F} \simeq \mathbf{Sh}_{\mathcal{E}}(Y)$, by 3.2.8. □

We remark that, for a closed inclusion f, the direct image functor f_* is regular (i.e. preserves epimorphisms) by A4.5.4, and hence preserves directed internal posets. So the elementary argument which we rejected in the proof of (iii) above would in fact suffice to prove the following corollary.

Corollary 3.2.17

(i) *A geometric morphism is proper iff both halves of its surjection–inclusion factorization are proper, iff it can be factorized as a proper surjection followed by a closed inclusion.*

(ii) *Proper geometric morphisms are closed.*

Proof (i) The first assertion is immediate from 3.2.16; the second follows from it and the fact that an inclusion is proper iff it is closed (cf. 3.2.9).

(ii) It is immediate from (i) (and the propriety of closed inclusions) that if $f : \mathcal{F} \to \mathcal{E}$ is proper then the image under f of any closed subtopos of $\mathcal{F}$ is a closed subtopos of $\mathcal{E}$. But the definition of propriety is stable under slicing, so this is sufficient. □

In seeking a characterization of propriety in terms of sites, we shall make the standing assumption that the underlying category of our site is coherent (in fact, it might as well be a pretopos, although we shall not assume this), and that its coverage contains the coherent coverage P of 2.1.12(d). It turns out that, in addition to the sieves in the coherent coverage, we need only consider those covers which are generated by directed families of monomorphisms: we shall say that a sieve R on an object U of a coherent category $\mathcal{C}$ is a *dm-sieve* if, whenever it contains a morphism $a : V \to U$, it also contains the monic part $\mathrm{im}\, a$ of a, and in addition the monomorphisms in R form a directed subset of $\mathrm{Sub}_{\mathcal{C}}(U)$. (Equivalently we could say that, for any finite subset R' of R, there is a monomorphism $m : U' \rightarrowtail U$ in R such that every morphism in R' factors through m.) And we call a coverage T a *dm-coverage* if it consists of *dm*-sieves.

We shall also require a 'compatibility condition' between the *dm*-coverage and the coherent coverage: we shall say that T is *P-compatible* if it satisfies the following two conditions:

(i) Given $R \in T(U)$ and a cover $a: U \twoheadrightarrow V$ in $\mathcal{C}$, the sieve on V generated by the images of the composites ab, $b \in R$, is T-covering.

(ii) Given monomorphisms $U \rightarrowtail W$, $V \rightarrowtail W$ with $W = U \cup V$, and a sieve $R \in T(U)$, the sieve on W generated by the subobjects $(\mathrm{im}\, a \cup V)$, $a \in R$, is T-covering.

Lemma 3.2.18 *Let $\mathcal{C}$ be a (small) coherent category.*

(i) *If J is a Grothendieck coverage on $\mathcal{C}$ containing the coherent coverage P, then the set J_d of all dm-sieves in J is a P-compatible Grothendieck coverage, and J is the join of P and J_d in the poset of Grothendieck coverages on $\mathcal{C}$.*

(ii) *Conversely, if D is a P-compatible Grothendieck dm-coverage on $\mathcal{C}$, then the join $P \vee D(U)$ consists of all sieves R on U for which there exists $S \in D(U)$ such that $a^*(R) \in P(V)$ for each $a: V \to U$ in S. In particular, D consists precisely of the dm-sieves in $P \vee D$.*

Proof (i) It is clear that J_d inherits the closure properties of a Grothendieck coverage from J. For the P-compatibility of J_d, we use the local character of J; for each of the sieves asserted by the compatibility conditions to lie in J_d, we see that there is a P-covering sieve (that generated by the single morphism a in (i), and that generated by $U \rightarrowtail W$ and $V \rightarrowtail W$ in (ii)) such that the pullback of the given sieve along each morphism in the P-covering sieve is J-covering. Finally, to show that J is the join of P and J_d, let R be an arbitrary J-covering sieve on U. Let S be the sieve on U generated by all monomorphisms $U' \rightarrowtail U$ which are the unions of the images of finite subsets of R. Then S is a *dm*-sieve (and in J_d, since it contains R), and the pullback of R along any member of S is P-covering. Hence any Grothendieck coverage containing both P and J_d must contain J; but J contains P by assumption and J_d by definition, so it must be their join.

(ii) Clearly, the join $P \vee D$ must contain all the sieves as described; to show that it contains only those sieves, we must show that they form a Grothendieck coverage – specifically, that they satisfy the local character axiom (L). For this, it suffices (since P and D separately satisfy (L)) to show that we can 'commute a D-cover past a P-cover': that is, if R is a sieve on U, and we can find $S \in P(U)$ such that $a^*(R) \in D(V)$ for each $a: V \to U$ in S, then we can also find $S' \in D(U)$ such that $b^*R \in P(W)$ for each $b: W \to U$ in S'. So suppose S contains a finite family $(a_i: V_i \to U \mid 1 \leq i \leq n)$, such that the union of the images of the a_i is the whole of U. We define S' to be the sieve generated by those monomorphisms $U' \rightarrowtail U$ for which there exist $(V_i' \rightarrowtail V_i) \in a_i^*R$ for each i such that U' is the

union of the images of the $V_i' \rightarrowtail V_i \to U$ $(1 \leq i \leq n)$. Clearly, S' is a *dm*-sieve; and the P-compatibility conditions (plus the local character of D, which ensures as we saw after 2.1.8 that $D(U)$ is closed under finite intersections) imply that $S' \in D(U)$. But the pullback of R along each $U' \rightarrowtail U \in S(U)$ contains a P-covering family $(V_i' \to U' \mid 1 \leq i \leq n)$, as required.

For the final assertion, suppose R is a *dm*-sieve on U and belongs to $P \vee D$. By the result just proved, there exists $S \in D(U)$ such that the pullback of R along each morphism $a\colon V \to U$ in S contains a finite P-covering family; but a^*R is a *dm*-sieve since R is, and hence it must contain 1_V. So $R \supseteq S$, and hence $R \in D(U)$. □

Theorem 3.2.19 *For a Grothendieck topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is compact, i.e. the geometric morphism $\mathcal{E} \to$ **Set** is proper.*

(ii) *There exists a small site of definition $(\mathcal{C}, T)$ for $\mathcal{E}$ such that $\mathcal{C}$ is a coherent category and T is the union of the coherent coverage P and a P-compatible dm-coverage D such that the terminal object of $\mathcal{C}$ is D-irreducible.*

(iii) *There exists a small site of definition $(\mathcal{C}, J)$ for $\mathcal{E}$ such that $\mathcal{C}$ is a coherent category and J is the join, in the poset of Grothendieck coverages on $\mathcal{C}$, of the coherent coverage P and a P-compatible dm-coverage D such that the terminal object of $\mathcal{C}$ is D-irreducible.*

(iv) *There exists a site $(\mathcal{C}, J)$ as in* (iii), *but with the added assumption that $\mathcal{C}$ is a pretopos.*

Proof (i) ⇒ (iv): Take $\mathcal{C}$ to be a small full subcategory of $\mathcal{E}$ containing a generating set, and closed under finite limits and colimits; then $\mathcal{C}$ is a pretopos by A2.4.5. Let J be the Grothendieck coverage on $\mathcal{C}$ induced by the inclusion $\mathcal{C} \to \mathcal{E}$; then J contains the coherent coverage since the inclusion is a coherent functor, and so it has a join decomposition of the required sort by 3.2.18(i). Finally, the compactness of $\mathcal{E}$ implies that the terminal object has no nontrivial J-covering *dm*-sieves.

(iv) ⇒ (iii) and (iii) ⇒ (ii) are trivial.

(ii) ⇒ (iii): Given $T = P \cup D$ as in (ii), let $\tilde{D}$ be the Grothendieck coverage generated by D as in 2.1.9, and let J be the join of P and $\tilde{D}$ in the poset of Grothendieck coverages on $\mathcal{C}$. By 2.4.10, $1_{\mathcal{C}}$ is still $\tilde{D}$-irreducible; also, the J-sheaves on $\mathcal{C}$ are exactly the functors which are sheaves for both P and $\tilde{D}$ (equivalently, for both P and D), so they coincide with the T-sheaves.

(iii) ⇒ (i): Clearly, any subterminal object of $\mathbf{Sh}(\mathcal{C}, J)$ is a union of subobjects of the form $l(U)$, where U is subterminal in $\mathcal{C}$ (and $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ is the canonical functor). So any covering of 1 by subterminal objects of $\mathbf{Sh}(\mathcal{C}, J)$ is refined by a J-covering of 1 by subterminal objects of $\mathcal{C}$; but 3.2.18(ii), plus the D-irreducibility of 1 in $\mathcal{C}$, implies that any covering of the latter sort has a finite subcover. So $\mathcal{E}$ is compact. □

The arguments of 3.2.18 and 3.2.19 are all constructive, and so may be applied to internal sites in any topos with a natural number object. As in one of the proofs of 2.4.11, the natural number object is needed for the proof of 3.2.19(i) $\Rightarrow$ (iv), in order to 'close off' the site under finite limits and colimits. However, in the case when $\mathcal{E}$ is localic over the base topos, we can dispense with the natural number object provided we sacrifice condition (iv) of 3.2.19; in this case we simply take C to be the (internal) poset of all subterminal objects of $\mathcal{E}$, which is a distributive lattice and hence a coherent category.

Before proceeding to the proof of the pullback-stability theorem, we note a result which extends 3.2.19 in the same way that 3.1.20 extended 3.1.19:

Scholium 3.2.20 *Let $p\colon \mathcal{E} \to \mathcal{S}$ be a bounded proper map, and let $(\mathbb{C}, J)$ be an internal site for $\mathcal{E}$ over $\mathcal{S}$ as in* 3.2.19(ii), (iii) *or* (iv). *Then p is surjective iff $\mathbb{C}$ is non-degenerate (i.e. the morphism $0 \to 1$ in $\mathbb{C}$ is non-invertible).*

Proof At first sight, it is rather surprising that the 'positive' condition of surjectivity should be (constructively) equivalent to a negative condition (non-degeneracy) on the site. However, the key point is that any finite object of $\mathcal{S}$ (in the Kuratowskian sense) is either inhabited or empty (cf. D5.4.5(ii)); so the same is true for P-covering sieves, and so the assertion that $0 \to 1$ is not invertible is equivalent to saying that every P-covering sieve on 1 is inhabited. Since the same is clearly true for D-covering sieves, the result follows from 2.4.8. □

Another explanation of why the 'negative' condition of 3.2.20 suffices for surjectivity is the fact that proper maps are closed (3.2.17(ii)): we should expect the negative condition to imply that p has dense image, but we already know its image is closed.

In theory, it should be possible to translate our site condition for propriety into a condition on a fibration of sites for it to induce a proper map of toposes, and hence to prove a stability theorem for proper maps under cofiltered limits, similar to 3.1.22 for open surjections. However, we shall not do this here; having already proved the stability of proper localic maps under cofiltered limits in 3.2.11, we shall be able to deduce the corresponding result for proper maps of toposes via the stability of the hyperconnected–localic factorization in 5.1.13 below. Instead, we now proceed with the proof of the pullback-stability theorem.

Theorem 3.2.21 *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$ where f is proper, and either g is bounded, or f is bounded and $\mathcal{E}$ has a natural number object. Then k is proper. If f is also surjective, then so is k.

Proof As in the 'open' case, this theorem has two entirely different proofs depending on which of f or g we assume to be bounded.

(i) For the case where f is bounded, we use our site characterization (3.2.19) for bounded proper maps: if $\mathcal{E}$ has a natural number object, we represent $\mathcal{F}$ as $\mathbf{Sh}_{\mathcal{E}}(\mathbb{C}, T)$ for an internal site satisfying the hypotheses of 3.1.14(ii), so that T is the union of the coherent coverage P and a P-compatible dm-coverage D. It is clear from the construction of $g^{\#}T$ that it is then the union of $g^{\#}P$ and $g^{\#}D$; but $g^{\#}P$ is simply the coherent coverage on $g^*\mathbb{C}$ by 2.4.7. It is straightforward to verify that $g^{\#}D$ inherits the properties of being a dm-coverage and P-compatible from D; also that the condition that the terminal object of $\mathbb{C}$ is D-irreducible 'lifts' to the corresponding condition for $g^*\mathbb{C}$. So $(g^*\mathbb{C}, g^{\#}T)$ is an internal site for $\mathcal{H}$ over $\mathcal{G}$ satisfying the hypotheses of 3.2.19(ii).

To deal with the 'surjective' case, we have simply to observe that the condition of 3.2.20 is preserved under the passage from $\mathbb{C}$ to $g^*\mathbb{C}$, since it is equivalent to saying that there is a pullback square

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & C_1 \arrow[d, "{(d_0, d_1)}"] \\ 1 \arrow[r, "{(\ulcorner 0 \urcorner, \ulcorner 1 \urcorner)}"] & C_0 \times C_0 \end{tikzcd} .$$

(ii) To show that arbitrary proper maps/surjections are stable under bounded pullback, it suffices as before to consider separately the cases when g has the form $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$ for an internal category $\mathbb{C}$, and when g is an inclusion. In fact, in place of the second of these, we shall consider the more general case when g is localic.

First suppose that $\mathcal{G}$ is of the form $[\mathbb{C}, \mathcal{E}]$ for an internal category $\mathbb{C}$ in $\mathcal{E}$. Since every slice of $\mathcal{G}$ is again of this form, it suffices to show that the square

$$\begin{tikzcd} {[f^*\mathbb{C}, \mathcal{F}]} \arrow[r, "{[\mathbb{C}, f]}"] \arrow[d, "\infty"] & {[\mathbb{C}, \mathcal{E}]} \arrow[d, "\infty"] \\ {[f^*\mathbb{I}, [f^*\mathbb{C}, \mathcal{F}]]} \arrow[r] & {[\mathbb{I}, [\mathbb{C}, \mathcal{E}]]} \end{tikzcd}$$

satisfies the Beck–Chevalley condition for subterminal objects, whenever $\mathbb{I}$ is a directed internal poset in $[\mathbb{C}, \mathcal{E}]$. But since the forgetful functor $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}/C_0$

is regular, it maps $\mathbb{I}$ to a directed internal poset ($\mathbb{J}$, say) in $\mathcal{E}/C_0$. Now consider the diagram

$$\begin{tikzcd}
{[f^*\mathbb{I},[f^*\mathbb{C},\mathcal{F}]]} \arrow[rr] \arrow[dr] \arrow[dd] & & {[\mathbb{I},[\mathbb{C},\mathcal{E}]]} \arrow[dr] \arrow[dd] & \\
& {[f^*\mathbb{C},\mathcal{F}]} \arrow[rr] \arrow[dd] & & {[\mathbb{C},\mathcal{E}]} \arrow[dd] \\
{[f^*\mathbb{J},\mathcal{F}/f^*C_0]} \arrow[rr] \arrow[dr] & & {[\mathbb{J},\mathcal{E}/C_0]} \arrow[dr] & \\
& \mathcal{F}/f^*C_0 \arrow[rr] & & \mathcal{E}/C_0
\end{tikzcd}$$

where the horizontal maps are all induced by f_*, the diagonal maps are colimit functors and the vertical maps are forgetful functors. It is easily verified that each of the vertical faces commutes up to isomorphism, and the forgetful functors reflect isomorphisms; so, from the fact that the bottom face commutes up to isomorphism at subterminal objects, we deduce that the top face does so too. The second assertion of the theorem is immediate, since surjectivity is always preserved under pullback to a topos of the form $[\mathbb{C},\mathcal{E}]$, as we saw in 3.1.24.

In the case where $g\colon \mathcal{G} \to \mathcal{E}$ is localic, we may form the hyperconnected–localic factorization of f and pull back each factor separately. The pullback of the hyperconnected part is hyperconnected by 2.4.11, and hence proper and surjective; so we are left with the case when f as well as g is localic. But if we translate the problem across the equivalence between localic $\mathcal{E}$-toposes and internal locales in $\mathcal{E}$, then the result we require is contained in 3.2.6. □

Corollary 3.2.22 *A geometric morphism is stably right weak Beck–Chevalley (in the sense of* 2.4.16*) iff it is proper.*

Proof One direction was established in 3.2.13. For the converse, since proper maps are stable under pullback it suffices to show that they are RWBC. But this may be proved in exactly the same way as 3.1.28: given a pullback of a proper morphism along a bounded morphism, we may use the hyperconnected–localic factorization of each side to factor the pullback square into four quarters, three of which have hyperconnected maps on a pair of opposite sides and so satisfy the weak Beck–Chevalley condition by A4.6.8. So we are reduced to the localic case; but in this case the result we require is contained in 3.2.6. □

As an application of the Beck–Chevalley characterizations of open and proper maps, we prove the following result, which will be of importance in Section C5.1.

Proposition 3.2.23 *Let*

$$\begin{tikzcd}
\mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

be a pullback square in $\mathfrak{BTop}$ *such that* f *is surjective and either open or proper, and* h *is either open or proper. Then* g *satisfies the same condition as* h.

Proof The proposition has four separate cases. The two when f and h are both open, or both proper, are easy; for in these cases the diagonal of the square is open (resp. proper) by 3.1.4(i) (resp. 3.2.16(i)), but k is surjective by 3.1.27 (resp. 3.2.21), and so g is open (resp. proper) by 3.1.4(ii) (resp. 3.2.16(ii)).

In the two cases when f and h satisfy opposite conditions, we exploit the Beck–Chevalley characterizations of open and proper maps (3.1.28 and 3.2.22). The two cases are very similar; we shall prove the one where h is open and f is proper, by showing that g is a SLWBC morphism. Suppose given a bounded morphism $p\colon \mathcal{E}' \to \mathcal{E}$; we may form the cube

$$\begin{tikzcd}
\mathcal{H}' \arrow[rr, "k'"] \arrow[dr, "s"] \arrow[dd, "h'"'] & & \mathcal{G}' \arrow[dr, "r"] \arrow[dd, "g'"' near end] & \\
& \mathcal{H} \arrow[rr, "k" near end] \arrow[dd, "h" near end] & & \mathcal{G} \arrow[dd, "g"] \\
\mathcal{F}' \arrow[rr, "f'" near end] \arrow[dr, "q"'] & & \mathcal{E}' \arrow[dr, "p"] & \\
& \mathcal{F} \arrow[rr, "f"] & & \mathcal{E}
\end{tikzcd}$$

in which all faces are pullbacks. Now the diagram

$$\begin{tikzcd} g^* f_* q_* \cong g^* p_* f'_* \arrow[r] \arrow[d, tail] & r_* g'^* f'_* \arrow[d, tail] \\ k_* h^* q_* \arrow[r, tail] & k_* s_* h'^* \cong r_* k'_* h'^* \end{tikzcd}$$

of Beck–Chevalley transformations commutes, and the indicated edges are monic because f and f' are proper and h is open. So $g^* p_* f'_*(B) \to r_* g'^* f'_*(B)$ is monic for any $B \in \text{ob}\,\mathcal{F}'$; but f' is surjective, since it is a pullback of f, and so for any $A \in \text{ob}\,\mathcal{E}'$ the unit map $A \to f'_* f'^*(A)$ is monic. Hence $g^* p_*(A) \to r_* g'^*(A)$ is monic for any A. Thus we have shown that g is a LWBC morphism; but the same argument may be applied to any pullback of g, so it is a SLWBC morphism.

In the other case, we similarly prove that $f'^* p^* g_* \to f'^* g'_* r^*$ is monic, and then use the fact that f'^* reflects monomorphisms. □

We have not had very much to say in this section about separated maps of toposes, as defined in 3.2.12(b). However, we may now note one interesting example:

Example 3.2.24 For a (discrete) group G, the topos $[G, \mathbf{Set}]$ is Hausdorff (that is, $[G, \mathbf{Set}] \to \mathbf{Set}$ is separated) iff G is finite. For we have an open surjection $s\colon \mathbf{Set} \to [G, \mathbf{Set}]$ whose inverse image is the forgetful functor; and if we pull back this morphism against itself, we obtain $\mathbf{Set}/G$ (as may easily be seen by identifying s with the morphism $[G, \mathbf{Set}]/G \to [G, \mathbf{Set}]$, where G acts on itself by left translation). In other words, the pullback of the diagonal of $[G, \mathbf{Set}]$ along $(s, s)\colon \mathbf{Set} \to [G, \mathbf{Set}] \times [G, \mathbf{Set}]$ is $\mathbf{Set}/G \to \mathbf{Set}$, which is manifestly proper iff G is finite. But since (s, s) is an open surjection, 3.2.23 implies that $\mathbf{Set}/G \to \mathbf{Set}$ is proper iff the diagonal of $[G, \mathbf{Set}]$ is proper.

For the record, we also note the basic stability properties of separated morphisms:

Lemma 3.2.25

(i) *Composites of separated morphisms are separated.*

(ii) *Pullbacks of separated morphisms are separated.*

(iii) *If fg is separated and g is a proper surjection, then f is separated.*

(iv) *If fg is separated and f is localic, then g is separated.*

Proof (i) Suppose $f\colon \mathcal{F} \to \mathcal{E}$ and $g\colon \mathcal{G} \to \mathcal{F}$ are separated. Then the diagonal of fg can be factored as the composite

$$\mathcal{G} \xrightarrow{\Delta_g} \mathcal{G} \times_{\mathcal{F}} \mathcal{G} \longrightarrow \mathcal{G} \times_{\mathcal{E}} \mathcal{G},$$

where the second factor is the pullback of Δ_f along $g \times g$; so the result follows from the stability of proper maps under composition and pullback.

(ii) is immediate from the stability of proper maps under pullback.

(iii) The composite $(\Delta_f)g$ is equal to $(g \times g)\Delta_{fg}$, which is proper by the hypotheses, so the result follows from 3.2.16(ii).

(iv) If f is localic then Δ_f is an inclusion by B3.3.8(ii); hence so is the second factor of the composite displayed in the proof of (i), and the result follows from 2.2.18(iii). □

Note in particular that a morphism is separated iff both halves of its hyperconnected–localic factorization are separated, iff the surjective part of its surjection–inclusion factorization is separated. (Inclusions are always separated, by B3.3.8(iii).)

Finally in this section, we return to the problem with which we began it:

Lemma 3.2.26 *A geometric morphism f is proper iff it is stably closed (i.e. all pullbacks of f along bounded morphisms are closed).*

Proof One direction is immediate from 3.2.21 and 3.2.17(ii). For the converse, we note that f is closed iff the localic part of its hyperconnected–localic factorization is closed, since a hyperconnected morphism $g: \mathcal{F} \to \mathcal{G}$ induces a bijection between closed subtoposes of $\mathcal{F}$ and of $\mathcal{G}$. And the hyperconnected–localic factorization is stable under pullback, by 2.4.12; so if f is stably closed, its localic part corresponds to a stably closed map of locales, and is therefore proper by 3.2.8. Hence f itself is proper. □

Suggestions for further reading: Moerdijk & Vermeulen [858], Plewe [979], Vermeulen [1203].

C3.3 Locally connected morphisms

It will be recalled that in A4.1.5 we defined a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ to be essential if its inverse image functor f^* has a left adjoint $f_!$ as well as its right adjoint f_*. However, the class of essential geometric morphisms is not as well-behaved as one might like. One reason (which we noted in B1.2.4) is that although, when we make $\mathcal{F}$ and $\mathcal{E}$ into $\mathcal{E}$-indexed categories, the adjunction $(f^* \dashv f_*)$ is always indexed over $\mathcal{E}$, the adjunction $(f_! \dashv f^*)$ need not be, even though $(f/A)^*$ has a left adjoint for every object A of $\mathcal{E}$. Accordingly, it is advantageous to study the more restricted class of morphisms $f: \mathcal{F} \to \mathcal{E}$ for which f^* has an $\mathcal{E}$-indexed left adjoint: for reasons adumbrated after 1.5.9 above, we call such morphisms *locally connected*.

Our first result in this section is the analogue over an arbitrary base of 1.5.9:

Proposition 3.3.1 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is locally connected.*

(ii) *'f^* commutes with Π-functors'; that is, for every morphism $h\colon A \to B$ in $\mathcal{E}$, the canonical (Beck–Chevalley) natural transformation in the square*

$$\begin{tikzcd} \mathcal{E}/A \arrow[r, "\Pi_h"] \arrow[d, "(f/A)^*"'] \arrow[dr, phantom, "\downarrow"] & \mathcal{E}/B \arrow[d, "(f/B)^*"] \\ \mathcal{F}/f^*A \arrow[r, "\Pi_{f^*h}"] & \mathcal{F}/f^*B \end{tikzcd}$$

is an isomorphism.

(iii) *For each object A of $\mathcal{E}$, the functor $(f/A)^*\colon \mathcal{E}/A \to \mathcal{F}/f^*A$ is cartesian closed.*

Proof (i) $\Leftrightarrow$ (ii): Since f^* always preserves finite limits, condition (ii) is just the assertion that f^* is continuous as an $\mathcal{E}$-indexed functor $\mathbb{E} \to \mathbb{F}$ (cf. B1.4.13). Since both these $\mathcal{E}$-indexed categories are complete and locally small, and $\mathbb{E}$ is well-powered and has a coseparator by B3.1.13, the result follows from the Indexed Adjoint Functor Theorem B2.4.6.

(ii) $\Rightarrow$ (iii): Given an object $h\colon B \to A$ of $\mathcal{E}/A$, consider the diagram

$$\begin{tikzcd} \mathcal{E}/A \arrow[r, "h^*"] \arrow[d, "(f/A)^*"] & \mathcal{E}/B \arrow[r, "\Pi_h"] \arrow[d, "(f/B)^*"] & \mathcal{E}/A \arrow[d, "(f/A)^*"] \\ \mathcal{F}/f^*A \arrow[r, "(f^*h)^*"] & \mathcal{F}/f^*B \arrow[r, "\Pi_{f^*h}"] & \mathcal{F}/f^*A \end{tikzcd}$$

of which the left-hand cell commutes since f^* preserves pullbacks, and the right-hand cell commutes by (ii). But the top and bottom composites are the functors $(-)^h$ and $(-)^{f^*h}$ respectively.

(iii) $\Rightarrow$ (ii) similarly follows from the construction of Π-functors from exponentials and pullbacks, given in the proof of A1.5.2(i). □

If we already know that f is essential, then the implication (iii) $\Rightarrow$ (i) of 3.3.1 may also be deduced from A1.5.8, as we observed in B1.2.4.

Corollary 3.3.2

(i) *Any stable left Beck–Chevalley morphism (in the sense of 2.4.16) is locally connected.*

(ii) *Locally connected morphisms are open.*

(iii) *A composite of locally connected morphisms is locally connected.*

(iv) *Given a commutative diagram*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

where h is locally connected and g is either connected or an open surjection, then f is locally connected.

Proof (i) is immediate from condition (ii) of 3.3.1.

(ii) If f^* commutes with the functors Π_h, then it commutes with the functors $\forall_h$ which are their restrictions to subterminal objects; so it is a Heyting functor. (Alternatively, we may use condition (iii) of 3.3.1 and (v) of 3.1.7.)

(iii) follows easily from (iii) of 3.3.1, since a composite of cartesian closed functors is cartesian closed.

(iv) In the case where g is connected, we have $f^* \cong g_*g^*f^* \cong g_*h^*$, so it has an $\mathcal{E}$-indexed left adjoint $h_!g^*$.

In the case where g is an open surjection, we use 3.3.1(iii); since the hypotheses are stable under slicing by an object of $\mathcal{E}$, it is sufficient to show that f^* is cartesian closed. Consider the exponential comparison map $\theta_{A,B}\colon f^*(B^A) \to f^*B^{f^*A}$, where A and B are objects of $\mathcal{E}$. We have a commutative diagram

$$\begin{tikzcd} h^*(B^A) \arrow[r, "g^*(\theta_{A,B})"] \arrow[dr] & g^*(f^*B^{f^*A}) \arrow[d, tail] \\ & h^*B^{h^*A} \end{tikzcd}$$

where the diagonal arrow is an isomorphism since h^* is a cartesian closed functor, and the vertical arrow is monic by 3.1.7(iv) since g is open. It follows easily that $g^*(\theta_{A,B})$ is an isomorphism; but g^* reflects isomorphisms since g is surjective. □

Once we have proved that locally connected morphisms are stable under pullback, we shall be able to establish the converse of 3.3.2(i).

We recall from 1.5.7 that f is said to be *connected* if f^* is full and faithful. For locally connected morphisms, there is a particularly simple criterion for

connectedness:

Lemma 3.3.3 *Let $f: \mathcal{F} \to \mathcal{E}$ be a locally connected geometric morphism. Then f is connected iff the left adjoint $f_!$ of f^* preserves* 1.

Proof If f is connected, then the counit of the adjunction $(f_! \dashv f^*)$ is an isomorphism, so $f_!(1) \cong f_! f^*(1) \cong 1$. The converse is immediate from A1.5.9(i), since f^* is cartesian closed. $\square$

We call a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ a *local homeomorphism* if there is an object A of $\mathcal{E}$ and an equivalence $\mathcal{F} \simeq \mathcal{E}/A$ identifying f^* and f_* with A^* and Π_A respectively (cf. A2.3.8). Clearly, a local homeomorphism is locally connected, since A^* is a logical functor (A2.3.2) and so commutes with Π-functors.

Lemma 3.3.4 *Connected morphisms are orthogonal to local homeomorphisms in $\mathfrak{Top}$; that is, given a commutative square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

where h is connected and g is a local homeomorphism, there exists (uniquely up to unique 2-isomorphism) a morphism $l: \mathcal{F} \to \mathcal{G}$ making both triangles commute.

Proof Suppose $\mathcal{G} \simeq \mathcal{E}/A$ as an $\mathcal{E}$-topos. Then by B3.2.8(b) the geometric morphism k corresponds to a morphism $1 \to h^* f^*(A)$ in $\mathcal{H}$; but h^* is full and faithful, so this derives from a unique morphism $1 \to f^*(A)$ in $\mathcal{F}$, which in turn corresponds to the required geometric morphism l. $\square$

Proposition 3.3.5

(i) *Any locally connected morphism can be factored, uniquely up to equivalence, as a connected locally connected morphism followed by a local homeomorphism.*

(ii) *A geometric morphism $f: \mathcal{F} \to \mathcal{E}$ is locally connected iff, for each object B of $\mathcal{F}$, the composite $\mathcal{F}/B \to \mathcal{F} \to \mathcal{E}$ may be factored as a connected morphism followed by a local homeomorphism.*

Proof (i) Given a locally connected morphism f, we define $A = f_!(1)$; then we certainly have a factorization $f_! = \Sigma_A g_!$, where $g_!(B) = f_!(B \to 1)$. And $g_!$ has a right adjoint g^* given by the composite

$$\mathcal{E}/A \xrightarrow{(f/A)^*} \mathcal{F}/f^*A \xrightarrow{\eta_1^*} \mathcal{F} ,$$

where $\eta_1 : 1 \to f^* f_!(1)$ is the unit map (cf. A4.1.13). But this composite is the inverse image of a locally connected geometric morphism, since each of its factors

is; and the resulting morphism $\mathcal{F} \to \mathcal{E}/A$ is connected by 3.3.3. This establishes the existence of the factorization; the uniqueness follows from 3.3.4.

(ii) If f is locally connected, then the existence of the factorization follows from (i) and 3.3.2(iii). Conversely, suppose the factorization exists for any $B \in \operatorname{ob} \mathcal{F}$; let us write $f_!B$ for the object of $\mathcal{E}$ yielding the factorization $\mathcal{F}/B \to \mathcal{E}/f_!B \to \mathcal{E}$ of $\mathcal{F}/B \to \mathcal{E}$. We claim first that $f_!$ defines a functor $\mathcal{F} \to \mathcal{E}$, left adjoint to f^*. For, if A is any object of $\mathcal{E}$, then morphisms $B \to f^*A$ in $\mathcal{F}$ correspond to (isomorphism classes of) geometric morphisms $\mathcal{F}/B \to \mathcal{F}/f^*A$ over $\mathcal{F}$, by B3.2.8(b); but these in turn correspond to geometric morphisms $\mathcal{E}/f_!B \to \mathcal{E}/A$ over $\mathcal{E}$, by 3.3.4, and hence to morphisms $f_!B \to A$ in $\mathcal{E}$.

Now the hypothesis that the factorizations exist is clearly 'stable under slicing'; so, to prove that f is locally connected, it suffices to show that f^* is cartesian closed, or equivalently (by A1.5.8) that the Frobenius reciprocity condition $f_!(B \times f^*A) \cong f_!B \times A$ holds. But it is easy to see that the induced morphism $\mathcal{F}/(B \times f^*A) \to \mathcal{E}/(f_!B \times A)$ is connected, since it is obtained by slicing $\mathcal{F}/B \to \mathcal{E}/f_!B$ over the object $(\pi_1 \colon f_!B \times A \to f_!B)$ of its codomain. So Frobenius reciprocity follows from the uniqueness of the factorization. □

We say an object A of (the domain of) a locally connected $\mathcal{S}$-topos $p \colon \mathcal{E} \to \mathcal{S}$ is *connected* if $p_!A \cong 1$; equivalently, if the composite geometric morphism $\mathcal{E}/A \to \mathcal{E} \to \mathcal{S}$ is connected as well as locally connected. From the fact that $p_!$ is an $\mathcal{S}$-indexed left adjoint, we may deduce what is really a reformulation of 3.3.5(ii):

Lemma 3.3.6 *If $p \colon \mathcal{E} \to \mathcal{S}$ is a locally connected topos over $\mathcal{S}$, then every object of $\mathcal{E}$ is expressible as an $\mathcal{S}$-indexed coproduct of connected objects.*

Proof Given $A \in \operatorname{ob} \mathcal{E}$, let $I = p_!A$, and let $\eta_A \colon A \to p^*I$ be the unit of the adjunction. Clearly, A is the I-indexed coproduct $\Sigma_I(\eta_A)$; we claim that $(p_!)^I$ maps η_A to the terminal object of $\mathcal{S}/I$, so that it may be regarded as an I-indexed family of connected objects. For this, we note that if $x \colon J \to I$ is any object of $\mathcal{S}/I$, then morphisms $\eta_A \to f^*x$ in $\mathcal{E}/f^*I$ correspond to morphisms $f_!A \to J$ in $\mathcal{S}$ which are sections of x, i.e. to morphisms $1_I \to x$ in $\mathcal{S}/I$. □

Thus we may think of $p_!A$, for a general object A of $\mathcal{E}$, as 'the object of connected components of A'. Saying that A is connected is equivalent to saying that it has no nontrivial '$\mathcal{E}$-indexed decompositions'; that is, any morphism $A \to p^*I$ factors through p^*x for some $x \colon 1 \to I$ in $\mathcal{S}$. Hence the decomposition of 3.3.6 is unique.

From either version of 3.3.2(iv), we know that the localic part of the hyperconnected–localic factorization of a locally connected morphism f is locally connected; it is even easier to see that it inherits connectedness from f. However, there seems to be no reason in general why the hyperconnected part of a locally connected morphism should be locally connected; we do not have a counterexample, but we digress briefly to mention an example of a hyperconnected morphism which is not locally connected.

Example 3.3.7 Consider the morphism $\gamma\colon \mathbf{Sh}(\mathcal{C}) \to \mathbf{Set}$, where $\mathcal{C}$ is the site derived from a small full subcategory of $\mathbf{Sp}$ as in A2.1.11(d). It is easy to see that $\mathbf{Sh}(\mathcal{C})$ has just two subterminal objects, and hence γ is hyperconnected by A4.6.6(v). However, if γ were essential, then for each object X of $\mathcal{C}$ the functor $A \mapsto \gamma^* A(X)$ would be representable (by $\gamma_! l(X)$, where $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C})$ is the Yoneda embedding); and it is easily verified that $\gamma^* A(X)$ is the set of locally constant functions $X \to A$, so that this functor fails to preserve infinite products if the space of components of X is not discrete. So, provided $\mathcal{C}$ contains a space whose space of components is not discrete, $\mathbf{Sh}(X)$ is not locally connected over $\mathbf{Set}$.

In the converse direction, a connected and locally connected morphism need not be hyperconnected (consider sheaves on a connected and locally connected space, such as $\mathbb{R}$); but we shall see in 3.5.4(i) that, if we replace local connectedness by the stronger condition of being atomic, then the implication does hold.

In A4.2.7(b) and 3.1.2, we gave necessary and sufficient conditions on a functor $f\colon \mathcal{C} \to \mathcal{D}$ between small categories for the corresponding geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ to be surjective (respectively open). It would clearly be convenient to have similar conditions for connectedness and local connectedness. In fact it seems hard to give conditions which are both necessary and sufficient; but there are sufficient conditions which are often useful. In order to state them, we introduce an auxiliary definition: given $f\colon \mathcal{C} \to \mathcal{D}$, we define Ext_f to be the category whose objects are quadruples (U, V, r, i) where $U \in \mathrm{ob}\,\mathcal{C}$, $V \in \mathrm{ob}\,\mathcal{D}$, $r\colon fU \to V$ and $i\colon V \to fU$ satisfy $ri = 1_V$, and whose morphisms $(U, V, r, i) \to (U', V', r', i')$ are pairs $(a\colon U \to U', b\colon V \to V')$ such that $r'(fa) = br$ and $i'b = (fa)i$. We shall write $g\colon \mathrm{Ext}_f \to \mathcal{D}$ for the obvious functor sending (U, V, r, i) to V, and $\mathrm{Ext}_f(V)$ for the fibre of this functor over a particular object V (we call this the category of f-*extracts* of V, the word 'extract' being used in this context as the opposite of 'retract').

We note that the conditions of A4.2.7(b) and 3.1.2 can both be formulated in terms of the notion of f-extract: the geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ induced by f is surjective iff $\mathrm{Ext}_f(V)$ is nonempty for each $V \in \mathrm{ob}\,\mathcal{D}$, and open iff g has the 'left lifting property' that, given $b\colon V \to V'$ in $\mathcal{D}$ and an object (U, V, r, i) of Ext_f mapped by g to the domain of b, we can find a morphism of Ext_f whose domain is (U, V, r, i) and which is mapped by g to b. (Actually, the condition for openness was stated in 3.1.2 as a special case of this, in which the domain of b was assumed to be of the form fU and its lifting to Ext_f was of the form $(U, fU, 1_{fU}, 1_{fU})$. However, the general case follows from this one: given $b\colon V \to V'$ and (U, V, r, i) as above, a lifting of the composite br in the sense of 3.1.2 is automatically a lifting of b in the sense just given.)

Lemma 3.3.8 *Let $f: \mathcal{C} \to \mathcal{D}$ be a functor between small categories.*

(i) *If $g: \mathrm{Ext}_f \to \mathcal{D}$ is surjective on morphisms and, for each $V \in \mathrm{ob}\,\mathcal{D}$, the category $\mathrm{Ext}_f(V)$ is connected, then the induced geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is connected.*

(ii) *Suppose that, for each $b: V \to V'$ in $\mathcal{D}$ and each lifting of V to an object (U, V, r, i) of Ext_f, the category of liftings of b (whose objects are liftings of b to a morphism of Ext_f with domain (U, V, r, i), and whose morphisms are morphisms of $\mathrm{Ext}_f(V')$ forming commutative triangles) is connected. Then $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is locally connected.*

Proof (i) Let $S, T: \mathcal{D} \rightrightarrows \mathbf{Set}$ be two functors and $\phi: f^*S \to f^*T$ be a natural transformation. We have to show that there is a (unique) natural transformation $\theta: S \to T$ with $f^*\theta = \phi$; but the uniqueness follows from A4.2.7(*b*), since the connectedness of $\mathrm{Ext}_f(V)$ implies that it is nonempty. To show existence, let V be any object of $\mathcal{D}$ and choose an object (U, V, r, i) of Ext_f lying over it; then we define θ_V to be the composite

$$SV \xrightarrow{Si} S(fU) \xrightarrow{\phi_U} T(fU) \xrightarrow{Tr} TV;$$

the connectedness of $\mathrm{Ext}_f(V)$ (plus naturality of ϕ) is exactly what we need to prove that this definition is independent of the choice of U, r and i. To verify that θ is natural, we need the other hypothesis: given $b: V \to V'$, let $(a, b): (U, V, r, i) \to (U', V', r', i')$ be a morphism of Ext_f lying over it, and consider the diagram

$$\begin{tikzcd}
SV \arrow[r, "Si"] \arrow[d, "Sb"] & S(fU) \arrow[r, "\phi_U"] \arrow[d, "S(fa)"] & T(fU) \arrow[r, "Tr"] \arrow[d, "T(fa)"] & TV \arrow[d, "Tb"] \\
SV' \arrow[r, "Si'"] & S(fU') \arrow[r, "\phi'_U"] & T(fU') \arrow[r, "Tr'"] & TV'
\end{tikzcd}$$

where the central square commutes by naturality of ϕ, and the other two by functoriality of S and T.

(ii) We note first that the hypothesis is stable under slicing: if we have a pullback square

$$\begin{tikzcd}
\mathcal{C}' \arrow[r, "f'"] \arrow[d, "d"] & \mathcal{D}' \arrow[d, "e"] \\
\mathcal{C} \arrow[r, "f"] & \mathcal{D}
\end{tikzcd}$$

where e (and hence also d) is a discrete opfibration, then f' inherits the condition from f. For in this case it is easy to see that a lifting of an object (U, V, r, i) of Ext_f to one of $\mathrm{Ext}_{f'}$ is uniquely specified by giving an element x of $T(V)$, where $T\colon \mathcal{D} \to \mathbf{Set}$ is the functor corresponding to e, since the element of $T(fU)$ corresponding to the lifting of U to an object of $\mathcal{C}'$ must be $Ti(x)$; and similarly for morphisms of $\mathrm{Ext}_{f'}$. (In other words, $\mathrm{Ext}_{f'} \to \mathrm{Ext}_f$ is also a discrete opfibration, and is simply the pullback of e along g.) Now, given a morphism $b\colon (V, x) \to (V', Tb(x))$ of $\mathcal{D}'$ and a lifting of its domain to $\mathrm{Ext}_{f'}$, the category of liftings of b to $\mathrm{Ext}_{f'}$ is easily seen to be isomorphic to the category of liftings of its image in $\mathcal{D}$ to Ext_f.

Hence, by 3.3.1(iii), it suffices to show that the hypothesis of (ii) implies that $f^*\colon [\mathcal{D}, \mathbf{Set}] \to [\mathcal{C}, \mathbf{Set}]$ is a cartesian closed functor. We recall the construction of exponentials in $[\mathcal{D}, \mathbf{Set}]$ from A1.5.5: given functors $S, T\colon \mathcal{D} \rightrightarrows \mathbf{Set}$ and $V \in \mathrm{ob}\, \mathcal{D}$, $T^S(V)$ is the set of natural transformations $\mathcal{D}(V, -) \times S \to T$. In terms of this description, it is easy to see that the comparison map $\theta_{S,T}\colon f^*(T^S) \to f^*T^{f^*S}$ sends a natural transformation $\alpha\colon \mathcal{D}(fU, -) \times S \to T$ to the transformation $\beta\colon \mathcal{C}(U, -) \times f^*S \to f^*T$ given by $\beta_{U'}(a\colon U \to U', x) = \alpha_{fU'}(fa, x)$. Thus we have to show that, given $\beta\colon \mathcal{C}(U, -) \times f^*S \to f^*T$, we can uniquely reconstruct α from it. But this is easy: given $b\colon fU \to V$ in $\mathcal{D}$ and $x \in S(V)$, we define $\alpha_V(b, x) = (Tr)\beta_{U'}(a, (Si)x)$, where $(a, b)\colon (U, fU, 1_{fU}, 1_{fU}) \to (U', V, r, i)$ is any morphism of Ext_f lying over b. The connectedness of the category of liftings of b ensures that this is independent of the choice of lifting; moreover, it is natural, since if we are given $b'\colon V \to V'$, then we may lift it to a morphism $(a', b')\colon (U', V, r, i) \to (U'', V', r', i')$ of Ext_f and we have

$$\begin{aligned}
(Tb')\alpha_V(b, x) &= (Tb')(Tr)\beta_{U'}(a, (Si)x) \\
&= (Tr')(T(fa'))\beta_{U'}(a, (Si)x) \\
&= (Tr')\beta_{U''}(a'a, S(fa')(Si)x) \\
&= (Tr')\beta_{U''}(a'a, (Si')(Sb')x) \\
&= \alpha_{V'}(b'b, (Sb')x)\ .
\end{aligned}$$

Finally, if b happens to be of the form fa, then we may take (a, b) as our lifting of it, and thus deduce that $(\theta_{S,T})_U(\alpha) = \beta$. □

We note that the arguments in the proof of 3.3.8 are constructive, and so can be applied to functors between internal categories in any topos $\mathcal{S}$. The same applies to the following example (which we shall meet again in 5.2.7), provided $\mathcal{S}$ has a natural number object.

Example 3.3.9 Let $\mathcal{D}$ be the category of finite cardinals (that is, sets of the form $[p] = \{0, 1, \ldots, p-1\}$ for some $p \in \mathbb{N}$; cf. A2.5.14) and arbitrary maps between them, and let $\mathcal{C}$ be the poset of finite partial equivalence relations

(that is, symmetric and transitive relations) on $\mathbb{N}$, ordered by inclusion. We have a functor $f\colon \mathcal{C} \to \mathcal{D}$ which sends a relation R to the cardinality of the set of R-equivalence classes (for definiteness, we suppose that the classes are listed in increasing order of their greatest members); clearly, if $R \subseteq S$, then each R-equivalence class is contained in a unique S-equivalence class, which explains how f is defined on morphisms. We claim that f satisfies the hypotheses of 3.3.8(i) and (ii), so that the geometric morphism $[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ which it induces is connected and locally connected.

To prove this, first suppose $b\colon [p] \to [q]$ is a morphism of $\mathcal{D}$. Let $[\![p]\!]$ denote the relation consisting of all pairs $\langle j, j\rangle$ with $j < p$; clearly, $f[\![p]\!] = [p]$. We define a larger relation R to consist of all pairs $\langle j, j\rangle$ with $j < p+q$, all pairs $\langle j, p+b(j)\rangle$ and $\langle p + b(j), j\rangle$ with $j < p$, and all pairs $\langle j, k\rangle$ where $j, k < p$ and $b(j) = b(k)$. It is then clear that $f(R) = [q]$ and that f maps the inclusion $[\![p]\!] \subseteq R$ to b; so we have shown that f is surjective on morphisms, from which it easily follows that $g\colon \mathrm{Ext}_f \to \mathcal{D}$ is also surjective.

To verify the other condition of 3.3.8(i), let $(R, [p], r, i)$ be an arbitrary object of $\mathrm{Ext}_f([p])$. We shall show that it can be connected in this category to the object $([\![p]\!], [p], 1_{[p]}, 1_{[p]})$. To do this, first choose a natural number n which is greater than any number involved in the relation R, and let R' be the relation obtained from R by adding all pairs $\langle j, j\rangle$ with $n \leq j < n+p$ and also all pairs $\langle n+j, k\rangle$ and $\langle k, n+j\rangle$ where $j < p$ and k is in the $i(j)$th equivalence class of R. Since i is injective, it is easy to see that R' is (symmetric and) transitive, and that the inclusion $R \subseteq R'$ lifts to a morphism $(R, [p], r, i) \to (R', [p], r', i')$ of $\mathrm{Ext}_f([p])$, where r' and i' are related to r and i by a suitable permutation of the set $f(R) = f(R')$. Now let S be the relation consisting of all pairs $\langle j, j\rangle$ with $n \leq j < n+p$; then the inclusion $S \subseteq R'$ is mapped by f to $i'\colon [p] \to f(R')$, and hence defines a morphism $(S, [p], 1_{[p]}, 1_{[p]}) \to (R', [p], r', i')$ in $\mathrm{Ext}_f([p])$. But we may connect the domain of this morphism to $([\![p]\!], [p], 1_{[p]}, 1_{[p]})$ by similarly enlarging S to a relation S' which contains all pairs $\langle j, n+j\rangle$ and $\langle n+j, j\rangle$ with $j < p$, and then shrinking down again to $[\![p]\!]$.

The verification of the condition of 3.3.8(ii) is similar, and we shall omit the details.

Next, we consider how local connectedness can be characterized in terms of sites. We saw in Section C3.1 that openness corresponds to the condition 'all covers inhabited'; in a similar way, local connectedness corresponds to 'all covers connected'. Given a sieve R on an object U of a category $\mathcal{C}$, we say R is *connected* if it is connected as a full subcategory of $\mathcal{C}/U$, i.e. if it is inhabited and any two morphisms $f_1, f_2 \in R$ can be linked by a zigzag of the form

$$\begin{tikzcd}
V_1 \arrow[r] \arrow[drr, "f_1"'] & V_3 \arrow[dr, "f_3"] & V_4 \arrow[l] \arrow[d, "f_4"] & \cdots & V_n \arrow[dll] & V_2 \arrow[l] \arrow[dlll, "f_2"] \\
 & & U & & &
\end{tikzcd}$$

with all the f_i in R. And we say a site $(\mathcal{C}, T)$ (where T is assumed to be sifted, but not necessarily a Grothendieck coverage) is *locally connected* if every cover in T is connected. As usual, it is clear that this definition makes sense for internal sites in an arbitrary topos $\mathcal{S}$.

Theorem 3.3.10 *Let $p: \mathcal{E} \to \mathcal{S}$ be a bounded geometric morphism. Then p is locally connected iff there exists a locally connected internal site $(\mathbb{C}, T)$ in $\mathcal{S}$ such that $\mathcal{E} \simeq \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ in $\mathfrak{Top}/\mathcal{S}$. Moreover, p is connected and locally connected iff there exists a locally connected site of definition for $\mathcal{E}$ whose underlying category $\mathbb{C}$ has a terminal object.*

Proof We shall, as usual, argue as if in the case $\mathcal{S} = \mathbf{Set}$; but we note that all the arguments used are constructive, and so can be internalized in an arbitrary base topos $\mathcal{S}$.

First suppose $\mathcal{E} = \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$ where $(\mathbb{C}, T)$ is a locally connected site. We claim first that every constant functor $\mathbb{C}^{\mathrm{op}} \to \mathcal{S}$ is a T-sheaf, i.e. that the inverse image functor $\mathbb{C}^*: \mathcal{S} \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ factors through the inclusion $\mathcal{E} \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ – the resulting factorization being the functor p^*, of course. (Strictly speaking, we ought to write $\mathbb{C}^{\mathrm{op}*}$ rather than $\mathbb{C}^*$; but we hope the reader will forgive this abuse of notation.) So consider the constant functor $\mathbb{C}^* A$ whose value is an object A of $\mathcal{S}$; to say that $\mathbb{C}^* A$ is a T-sheaf is to say that , for each T-covering sieve R on an object U of $\mathbb{C}$, the unit map $A \to \varprojlim_{\mathbb{R}} \mathbb{R}^* A$ is an isomorphism, where $\mathbb{R}$ is the full subcategory of $\mathbb{C}/U$ corresponding to R. But for connected $\mathbb{R}$ this follows from B2.5.6.

Now $\mathbb{C}^*$ has a left adjoint $\varinjlim_{\mathbb{C}}$, which is $\mathcal{S}$-indexed by (a special case of) B2.3.21; so we obtain the desired $\mathcal{S}$-indexed left adjoint for p^* simply by restricting this functor to the full subcategory $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, T)$. And if $\mathbb{C}$ has a terminal object 1, then the unit map $A \to p_* p^*(A)$ may be identified with the unit map $A \to \varprojlim_{\mathbb{C}} \mathbb{C}^* A \cong \mathbb{C}^* A(1)$, which is an isomorphism; so p is connected.

Conversely, suppose p is locally connected. Then, as we saw in 3.3.6, every object of $\mathcal{E}$ is expressible as an $\mathcal{S}$-indexed coproduct of connected objects; so, given a bound B for $\mathcal{E}$ over $\mathcal{S}$, the family of all connected subobjects of B is a separating family for $\mathcal{E}$ as an $\mathcal{S}$-indexed category, and we can use it to form an internal full subcategory $\mathbb{C}$ of $\mathcal{E}$ which, when equipped with the coverage induced from the canonical coverage on $\mathcal{E}$, will yield a site of definition for $\mathcal{E}$ over $\mathcal{S}$. We must show that this site is locally connected. Clearly, every cover of an object of $\mathbb{C}$ must be inhabited, since connected objects are positive. And if R is any cover of an object U of $\mathbb{C}$, we observe that if $f_1: V_1 \to U$ and $f_2: V_2 \to U$ are morphisms in R such that the pullback $V_1 \times_U V_2$ (in $\mathcal{E}$) is inhabited, then they must lie in the same connected component of R, for some connected subobject of B must map to the pullback. Hence, if we write $\{S_i \mid i \in I\}$ for the set of connected components of R, and U_i, for each $i \in I$, for the subobject (in $\mathcal{E}$) of U which is the union of the images of all the morphisms in S_i, then the U_i form a pairwise-disjoint covering of U by subobjects; that is, we have a well-defined morphism

$U \to p^*I$ in $\mathcal{E}$ which sends U_i to the element i. So by the remark after 3.3.6 one of the U_i must be the whole of U, and the corresponding S_i is the whole of R. And, once again, if p is connected as well as locally connected, then the site just described will contain the terminal object of $\mathcal{E}$, provided it occurs as a subobject of B, and we can easily ensure the latter. □

Examples 3.3.11 (*a*) If a category $\mathcal{C}$ satisfies the right Ore condition (cf. A2.1.11(*h*)), then any inhabited sieve on an object of $\mathcal{C}$ is connected. Hence we may conclude from 3.3.10 that, for such a $\mathcal{C}$, any dense subtopos of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is locally connected. (As stated, this example is dependent on classical logic in **Set**, since denseness of a subtopos of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is equivalent to saying that covers are nonempty rather than inhabited; but we can constructivize it by re-interpreting the word 'dense' in the strong 'fibrewise' sense introduced in 1.1.22.)

(*b*) It is clear that a sieve generated by a single morphism is always connected; hence the site obtained by equipping a regular category with its regular coverage (2.1.12(*b*)) satisfies the conditions of 3.3.10. In view of D3.1.4, this says that *the classifying topos of any regular theory is connected and locally connected.*

(*c*) We revert to an example that we have considered already (for the case $\mathcal{S} = \mathbf{Set}$) in A2.1.11(*g*) (and see also D4.1.9). In the present context, let us take an arbitrary object I of $\mathcal{S}$ (which is here assumed to have a natural number object); let $\mathbb{P}$ be the poset of all partial maps $f \colon N \rightharpoonup I$ with finite domain (which may be interpreted as meaning either that the domain of f *is* a finite cardinal, considered as a subobject of N, or that it is isomorphic to a finite cardinal; it does not make any essential difference), ordered by the relation 'is an extension of'. We impose a coverage on $\mathbb{P}$ by saying that, for each $f : P$ and each $i : I$, the set

$$\{g \colon P \mid (g \text{ extends } f) \text{ and } (i \in \operatorname{im} g)\}$$

covers f. In A2.1.11(*g*) we verified that all such covers are connected (in the case $\mathcal{S} = \mathbf{Set}$, but the arguments used there are constructive); and $\mathbb{P}$ clearly has a top element, namely the empty partial map. So from 3.3.10 we may deduce that $\mathbf{Sh}_{\mathcal{S}}(\mathbb{P}, T)$ is a connected and locally connected $\mathcal{S}$-topos. We note also that this topos is the classifying topos (over $\mathcal{S}$) for the theory of partial surjections $N \rightharpoonup I$; that is, the intersection of the propositional theories which we considered in 1.2.8 and 1.2.9.

As usual, we may convert (the connected case of) 3.3.10 into a characterization of connected locally connected morphisms in terms of fibrations of sites. In order to state the condition, we need an *ad hoc* definition: given a functor $F \colon \mathbb{D} \to \mathbb{C}$, a sieve S on an object V of $\mathbb{D}$ and a coverage J on $\mathbb{C}$, we shall say

that R is *J-locally connected* if, whenever we are given a commutative square

$$\begin{tikzcd} U \arrow[r, "a_2"] \arrow[d, "a_1"] & FV_2 \arrow[d, "Fb_2"] \\ FV_1 \arrow[r, "Fb_1"] & FV \end{tikzcd}$$

with both b_1 and b_2 in S, we can find a J-covering sieve R on U such that, for each $U' \to U$ in R, we can connect b_1 to b_2 in S by a zigzag of morphisms, and also find a cone over the image of this zigzag under F with vertex U', connecting the given morphisms $U' \to U \to FV_1$ and $U' \to U \to FV_2$.

Proposition 3.3.12 *Let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism between bounded $\mathcal{S}$-toposes. Then f is connected and locally connected iff it can be induced by a fibration of sites $(P,T): (\mathbb{D}, K) \to (\mathbb{C}, J)$ in $\mathcal{S}$ such that P preserves covers and every K-covering sieve is J-locally connected.*

Proof First suppose f is connected and locally connected. Choose any site $(\mathbb{C}, J)$ for $\mathcal{E}$ over $\mathcal{S}$; then choose a locally connected site $(\mathbb{B}, H)$ with a terminal object for $\mathcal{F}$ over $\mathcal{E}$. If we externalize the latter to a site $(\mathbb{D}, K)$ in $\mathcal{S}$ as in 2.5.7, then the condition (in the internal logic of $\mathcal{E}$) that every H-covering sieve is (inhabited and) connected translates straightforwardly into the given conditions on K.

Conversely, suppose f is induced by a continuous fibration satisfying the conditions. We claim that if A is any J-sheaf on $\mathbb{C}$, then the composite $A \circ P$ is a K-sheaf: to see this, suppose given a K-covering sieve S on an object V and a compatible family of elements $(x_b \in A(PV') \mid b: V' \to V \in S)$. We must show that this family can be extended to a compatible family for the (J-covering) sieve on PV generated by the Pb, $b \in S$. But this is precisely the statement that, if we are given any commutative square

$$\begin{tikzcd} U \arrow[r, "a_2"] \arrow[d, "a_1"] & FV_2 \arrow[d, "Fb_2"] \\ FV_1 \arrow[r, "Fb_1"] & FV \end{tikzcd}$$

with b_1 and b_2 in S, then $A(a_1)(x_{b_1}) = A(a_2)(x_{b_2})$; so the assertion that U can be covered by morphisms $U' \to U$ for which we have a zigzag as above (plus the information that A is a J-sheaf) is precisely what we need to verify this.

It follows that in this case the inverse image f^* is simply the functor 'compose with P' applied to J-sheaves; since we also know that f_* is given by composition

with T, and since $PT \cong 1_{\mathbb{C}}$, it follows that $f_* f^* \cong 1_{\mathcal{E}}$, i.e. f is connected. But since P is a fibration, it is also easy to see that it satisfies the dual of the condition of 3.3.8(ii) (for the category of liftings of a given morphism has a terminal object, namely the prone lifting); hence composition with P commutes with Π-functors as a functor $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}] \to [\mathbb{D}^{\mathrm{op}}, \mathcal{S}]$. But the inclusions $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J) \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ and $\mathbf{Sh}_{\mathcal{S}}(\mathbb{D}, K) \to [\mathbb{D}^{\mathrm{op}}, \mathcal{S}]$ are cartesian closed by A4.2.9, and hence also commute with Π-functors; hence $(-) \circ P$ commutes with Π-functors as a functor between the sheaf toposes, i.e. f is locally connected. □

Corollary 3.3.13 *Let*

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

be an inverse sequence in $\mathfrak{BTop}/\mathcal{S}$, where each f_i is connected and locally connected. Then the legs $f_\infty^n \colon \mathcal{E}_\infty \to \mathcal{E}_n$ of the limit cone in $\mathfrak{BTop}/\mathcal{S}$ are connected and locally connected.

Proof This is not as simple as it seems, since it is not immediately obvious that the condition of 3.3.12 is stable under composition of fibrations of sites. Once one has verified that it is (which is straightforward but tedious), it is then easy to see that the same condition is inherited by the limit fibrations of 2.5.11. We omit the details, which may be found in [828]. □

As an application of 3.3.10, we prove an interesting extension of A4.5.18, which will be needed in the next section.

Theorem 3.3.14 *Let $f \colon \mathcal{F} \to \mathcal{E}$ be a bounded geometric morphism, and j a local operator on $\mathcal{E}$. Let k be the local operator on $\mathcal{F}$ making*

$$\begin{tikzcd} \mathbf{sh}_k(\mathcal{F}) \arrow[r] \arrow[d] & \mathbf{sh}_j(\mathcal{E}) \arrow[d] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

a pullback (cf. A4.5.14(e)). Then the corresponding subobject $K \rightarrowtail \Omega_{\mathcal{F}}$ is simply the upward closure of the image of the composite

$$f^*J \rightarrowtail f^*\Omega_{\mathcal{E}} \xrightarrow{\phi_1} \Omega_{\mathcal{F}} .$$

Proof First we prove the result under the assumption that f is locally connected as well as bounded. It is clear that K contains the upward closure of the image of $f^*J \to \Omega_{\mathcal{F}}$, so it suffices to prove that the classifying map of the upward closure (K', say) is a local operator. Suppose $\mathcal{F} \simeq \mathbf{Sh}_{\mathcal{E}}(\mathbb{C}, T)$, where $(\mathbb{C}, T)$ is a locally connected site; then $f^*\Omega_{\mathcal{E}}$ and f^*J are simply the appropriate constant

diagrams on $\mathbb{C}^{\mathrm{op}}$, by the proof of 3.3.10. On the other hand, $\Omega_{\mathcal{F}}$ is the internal diagram corresponding to the functor $\mathbb{C}^{\mathrm{op}} \to \mathbb{E}$ which (in set-theoretic terms) sends an object U of $\mathbb{C}$ to the set of T-closed sieves on U; and the comparison map $(\phi_1)_U$ sends a truth-value $p: \Omega_{\mathcal{E}}$ to the sieve $\{a: V \to U \mid p\}$ (note that such sieves are indeed T-closed, since T-covering sieves are inhabited). Hence a T-closed sieve R belongs to $K'(U)$ iff there exists a T-covering sieve S and a J-dense truth-value p such that $(p \Rightarrow [\![a \in R]\!])$ holds for all $a \in S$; but this implies that the truth-value $[\![1_U \in R]\!]$ is j-dense. It now follows easily that a subobject $B' \rightarrowtail B$ in $\mathcal{F}$ belongs to the class $\Xi(K')$ (that is, its classifying map factors through $K' \rightarrowtail \Omega_{\mathcal{F}}$) iff the sentence

$$(\forall U: C_0)(\forall x: B(U))([\![x \in B'(U)]\!] \in \ulcorner J \urcorner)$$

is valid in $\mathcal{E}$. Using this, it is trivial to verify that $\Xi(K')$ satisfies the conditions of A4.5.11(iii), and hence that the clasifying map of K' is a local operator.

Now consider the general case. Let $\mathcal{G}$ be the topos obtained by glueing along the cartesian functor $f^*: \mathcal{E} \to \mathcal{F}$, that is the category whose objects are triples (A, B, α) with $\alpha: B \to f^*A$. Then we have a closed inclusion $h: \mathcal{F} \to \mathcal{G}$ given by $h^*(A, B, \alpha) = B$ and $h_*(B) = (1, B, B \to 1)$ (cf. A4.5.6); but it is easy to check that we also have a locally connected geometric morphism $g: \mathcal{G} \to \mathcal{E}$ defined by setting $g^*(A) = (A, f^*A, 1_{f^*A})$, $g_!(A, B, \alpha) = A$ and $g_*(A, B, \alpha)$ to be the pullback of

$$\begin{tikzcd} & f_*B \arrow[d, "f_*(\alpha)"] \\ A \arrow[r, "\eta_A"] & f_*f^*A \end{tikzcd}$$

where η is the unit of $(f^* \dashv f_*)$. Clearly, the composite gh is isomorphic to f. Also, g is bounded if f is: if G is a bound for $\mathcal{F}$ over $\mathcal{E}$, then the object $H = (1, G, G \to 1)$ may be verified to be a bound for $\mathcal{G}$ over $\mathcal{E}$, as follows. Given an arbitrary object (A, B, α), choose a diagram

$$B \overset{e}{\twoheadleftarrow} \overline{B} \overset{m}{\rightarrowtail} f^*C \times G$$

where we may assume (on replacing C by $C \amalg 1$, if necessary) that C is well-supported. Now we have a diagram

$$\begin{tikzcd} (A \times C, \overline{B}, (\alpha e, \pi_1 m)) \arrow[r, tail, "{(1,(\alpha e, m))}"] \arrow[d, two heads, "{(\pi_1, e)}"] & (A \times C, f^*(A \times C) \times G, \pi_1) \cong g^*(A \times C) \times H \\ (A, B, \alpha) & \end{tikzcd}$$

of the required form in $\mathcal{G}$.

Thus we know that the conclusion of the theorem holds for g, by the first part of the proof. But it also holds for h, by A4.5.18(ii); and it is easy to see that the conclusion, regarded as a property of geometric morphisms, is 'stable under composition', since inverse image functors preserve everything involved in constructing the upward closure of the image of a morphism into a poset. So the conclusion holds for f. □

Next, we consider stability of (local) connectedness under pullback. As previously for openness and propriety, there are two versions of the pullback-stability theorem, depending on which side of the pullback square we assume to be bounded. One version follows readily from 3.3.10:

Theorem 3.3.15 *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "h"] \arrow[d, "k"] & \mathcal{F} \arrow[d, "f"] \\ \mathcal{G} \arrow[r, "g"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$ *with* f *locally connected, and either* f *or* g *bounded. Then* k *is locally connected, and the Beck–Chevalley natural transformation* $f^*g_* \to h_*k^*$ *is an isomorphism. Moreover, if* f *is also connected, then so is* k.

Proof (i) For the case when f is bounded, we have to verify that the property of being a locally connected site is preserved under change of base along g, as we defined it in Section C2.4. This is not quite automatic, since connectedness of a sieve (unlike inhabitedness) is not a first-order property. However, it is definable in terms of a coequalizer diagram, and g^* preserves coequalizers; hence, if R is a connected sieve on an object U of an internal category $\mathbb{C}$ in $\mathcal{E}$, we see that g^*R is a connected sieve on $g^*(U)$ in $g^*\mathbb{C}$. It follows that, if T is a coverage on $\mathbb{C}$ all of whose sieves are connected, then $g^{\#}T$ also consists of connected sieves; so the pullback-stability result is established. And if $\mathbb{C}$ has a terminal object, then so does $g^*\mathbb{C}$; so the connectedness assertion is also immediate.

For the Beck–Chevalley condition, we consider first the case when f has the form $[\mathbb{C}^{\mathrm{op}}, \mathcal{E}] \to \mathcal{E}$; then $\mathcal{H}$ is the topos $[g^*\mathbb{C}^{\mathrm{op}}, \mathcal{G}]$, and h_* is the functor 'apply g_* to discrete fibrations over $g^*\mathbb{C}$, and then pull back along the unit map $\mathbb{C} \to g_*g^*\mathbb{C}$'. But if we apply this to a constant diagram $(g^*\mathbb{C})^*B$, where B is an object of $\mathcal{G}$, we obtain the discrete fibration $\mathbb{C} \times g_*B \to \mathbb{C}$; thus we see that $f^*g_* \to h_*k^*$ is an isomorphism in this case. In general, when $\mathcal{F}$ is a subtopos of $[\mathcal{C}^{\mathrm{op}}, \mathcal{E}]$, we have seen that f^* is simply the factorization of the 'constant diagram' functor through the subcategory $\mathbf{Sh}_{\mathcal{E}}(\mathbb{C}, T)$, and similarly for k^*; and h_* is similarly the restriction of the direct image functor $[g^*\mathbb{C}^{\mathrm{op}}, \mathcal{G}] \to [\mathbb{C}^{\mathrm{op}}, \mathcal{E}]$ to sheaves, by A4.3.12. So the Beck–Chevalley condition in the general case follows from the particular one already established.

(ii) For the second version of the theorem, we shall as usual deal separately with the cases when g has the form $[\mathbb{C},\mathcal{E}] \to \mathcal{E}$ for an internal category $\mathbb{C}$ in $\mathcal{E}$, and when g is an inclusion.

If g has the form $[\mathbb{C},\mathcal{E}] \to \mathcal{E}$, we recall from B2.3.12(a) that any $\mathcal{E}$-indexed category extends in a canonical way to a $\mathbf{Cat}(\mathcal{E})$-indexed one: this extension sends the canonical indexing $\mathbb{E}$ of $\mathcal{E}$ over itself to the pseudofunctor $\mathbb{C} \mapsto [\mathbb{C},\mathcal{E}]$, and the $\mathcal{E}$-indexed category $\mathbb{F}$ to $\mathbb{C} \mapsto [f^*\mathbb{C},\mathcal{F}]$. It is clear that $\mathcal{E}$-indexed functors and adjunctions similarly extend; so, from the fact that $f^*\colon \mathbb{E} \to \mathbb{F}$ has an $\mathcal{E}$-indexed left adjoint, we deduce that it has a $\mathbf{Cat}(\mathcal{E})$-indexed one. But if we now restrict the latter indexing to the full subcategory of $\mathbf{Cat}(\mathcal{E})/\mathbb{C}$ consisting of discrete opfibrations, we obtain a $[\mathbb{C},\mathcal{E}]$-indexed left adjoint for $[\mathbb{C},f]^* = k^*$. Also, the Beck–Chevalley condition in this case is equivalent to saying that the natural map $k_!h^* \to g^*f_!$ is an isomorphism; but this is just part of the assertion that $f_!$ is a $\mathbf{Cat}(\mathcal{E})$-indexed functor.

In the case when g is an inclusion, let j be the corresponding local operator on $\mathcal{E}$, and j' the pullback local operator on $\mathcal{F}$ (cf. A4.5.14(e)). We claim that in this case the functor f^* maps j-sheaves to j'-sheaves, so that k^* is simply the restriction of f^* to $\mathbf{sh}_j(\mathcal{E})$. For if A is a j-sheaf in $\mathcal{E}$, then by A4.4.10 the canonical morphism $\psi_A\colon J^*A \to \Pi_d A$ (the transpose of the isomorphism $d^*J^*A \cong A$, where $d\colon 1 \rightarrowtail J$ is the generic j-dense monomorphism) is an isomorphism in $\mathcal{E}/J$. Since f^* commutes with Π-functors, we deduce that we have an isomorphism

$$(f^*J)^*(f^*A) \longrightarrow \Pi_{f^*d}(f^*A)$$

in $\mathcal{F}/f^*J$. Unpacking this as in A4.4.10, we deduce that f^*A satisfies the sheaf condition for all pullbacks of $f^*d\colon 1 \rightarrowtail f^*J$ in $\mathcal{F}$; hence the largest local operator on $\mathcal{F}$ for which f^*A is a sheaf (cf. A4.5.15) has the property that f^*m is dense for it whenever m is a j-dense monomorphism in $\mathcal{E}$. But j' is by definition the *smallest* local operator with this property; so f^*A must be a sheaf for it. It now follows easily that $k^*\colon \mathbf{sh}_j(\mathcal{E}) \to \mathbf{sh}_{j'}(\mathcal{F})$ has a left adjoint, namely $f_!$ followed by the reflector $\mathcal{E} \to \mathbf{sh}_j(\mathcal{E})$; and the fact that this adjunction is $\mathbf{sh}_j(\mathcal{E})$-indexed follows readily from the fact that $(f_! \dashv f^*)$ is $\mathcal{E}$-indexed. And the Beck–Chevalley condition in this case is just the assertion that k^* coincides with the restriction of f^* to sheaves.

For the final assertion of the theorem, we note that the Beck–Chevalley condition transposes to yield an isomorphism $k_!h^* \cong g^*f_!$; hence if $f_!$ preserves 1 then so does $k_!$, and the result follows from 3.3.3. □

Corollary 3.3.16 *A geometric morphism is stably left Beck–Chevalley (in the sense of* 2.4.16*) iff it is locally connected.*

Proof One direction was proved in 3.3.2(i); the converse follows from 3.3.15. □

Example 3.3.17 In the absence of local connectedness, (global) connectedness of geometric morphisms is not in general preserved under pullback. For a

counterexample, consider the pullback diagram

$$\begin{tikzcd} T \arrow[r] \arrow[d] & X \arrow[d, "g"] \\ S \arrow[r, "f"] & Y \end{tikzcd}$$

in **Loc**, where X is the unit interval $[0,1]$with its usual topology, Y is the same set with the topology whose only nontrivial open sets are the intervals $(t,1]$ $(0 < t < 1)$, S is the subspace $\{0\} \cup (\frac{1}{2}, 1]$ of Y, f is the inclusion and g is the identity mapping. Since S is complemented as a sublocale of Y, being the union of an open and a closed sublocale, it follows from 1.2.13 that the pullback T in **Loc** coincides with the pullback in **Sob** – that is, it is simply S retopologized as a subspace of X. In particular, it is disconnected, but S is connected; hence the identity mapping $T \to S$ cannot induce a connected geometric morphism. However, we claim that the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by g is connected.

To see this, note first that if $F \to Y$ is any local homeomorphism, then its pullback $E \to X$ along g is simply obtained by retopologizing the space F with the appropriate finer topology. Moreover, if we regard F and Y as posets under the usual specialization ordering, then the projection $F \to Y$ is a discrete opfibration (since each point of F has a neighbourhood mapping bijectively onto an open set in Y); and the open sets in F are exactly those which are open in E and upper sets in the specialization ordering. Using these facts, it is easy to see that if $F' \to Y$ is another local homeomorphism and $h \colon F \to F'$ is a mapping over Y, then h is continuous as a mapping $F \to F'$ iff it is continuous as a mapping $E \to E'$. So g^* is full and faithful.

Finally in this section, we fulfil a promise made in Section B4.5, by characterizing the algebras for the 'measure topos' monad $\mathbb{M}$ on $\mathfrak{BTop}/\mathcal{S}$ by a linear (rather than pointwise) cocompleteness condition, in the sense of B1.1.16.

Proposition 3.3.18 *Let $\mathcal{S}$ be a topos with a natural number object. Then a bounded $\mathcal{S}$-topos is an $\mathbb{M}$-algebra iff it is linearly $\mathcal{B}$-cocomplete, where $\mathcal{B}$ is the class of locally connected morphisms between bounded $\mathcal{S}$-toposes.*

Proof In B4.5.13, we saw that a topos $\mathcal{E}$ is an $\mathbb{M}$-algebra iff it is pointwise cocomplete with respect to the class $\mathcal{A}$ of $\mathcal{S}$-essential geometric morphisms. By B1.1.18, this is equivalent to saying that $\mathcal{E}$ is linearly cocomplete with respect

to the class of morphisms which appear as the top edges of comma squares

$$\begin{tikzcd} (f \downarrow g) \arrow[r] \arrow[d] \arrow[dr, phantom, "\to"] & \mathcal{H} \arrow[d, "g"] \\ \mathcal{G} \arrow[r, "f"] & \mathcal{F} \end{tikzcd}$$

whose bottom edges are in $\mathcal{A}$. Since all such morphisms are locally connected by 2.3.17, it is immediate that 'linearly $\mathcal{B}$-cocomplete' implies 'pointwise $\mathcal{A}$-cocomplete'. Conversely, if $\mathcal{E}$ is an $\mathbb{M}$-algebra then it is certainly $\mathcal{B}$-cocomplete, since $\mathcal{B} \subseteq \mathcal{A}$; the linearity condition follows from the Beck–Chevalley condition for pullbacks of locally connected morphisms (3.3.16), just as the pointwise condition for $\mathcal{A}$-cocompleteness followed from that for comma squares in the proof of B4.5.13. □

Suggestions for further reading: Barr & Paré [84], Moerdijk [828].

C3.4 Tidy morphisms

As we noted after 1.5.5, the condition that the direct image functor $\mathbf{Sh}(X) \to \mathbf{Set}$ should preserve directed unions of subterminal objects, which is equivalent to compactness of the locale X, seems less 'natural' in topos-theoretic terms than the condition that this functor should preserve all filtered colimits. But the latter is appreciably stronger: we shall call a locale X (or more generally a **Set**-topos $\mathcal{E}$) *strongly compact* if the direct image functor $\gamma_* \colon \mathbf{Sh}(X) \to \mathbf{Set}$ (resp. $\gamma_* \colon \mathcal{E} \to \mathbf{Set}$) preserves filtered colimits.

Examples 3.4.1 (*a*) Let X be the topological space obtained from two copies of the closed unit interval $[0,1]$ by identifying the two copies of the open interval $(0,1)$. It is clear that X is compact, since $[0,1]$ is. But, for any $n > 0$, let X_n denote the space obtained from two copies of $[0,1]$ by identifying the two copies of the interval $(1/2^{n+1}, 1 - 1/2^{n+1})$: it is clear that we have surjective local homeomorphisms $X_n \to X_m$ whenever $n < m$, and that $X = X_\infty$ is the colimit of this diagram in **LH** (and hence in $\mathbf{LH}/X \simeq \mathbf{Sh}(X)$). So, considering the X_n as sheaves on X, they form a directed diagram whose colimit has a (unique) global section, but no X_n has a global section. Hence X is not strongly compact.

(*b*) For any group G, it is clear that the topos $[G, \mathbf{Set}]$ is compact, since it is hyperconnected over **Set**. But $[G, \mathbf{Set}]$ *is strongly compact iff G is finitely generated.* For, if G is not finitely generated, we can express it as a nontrivial directed union of subgroups $(G_i \mid i \in I)$; then the G-sets G/G_i form a directed diagram in $[G, \mathbf{Set}]$ whose colimit is the terminal object, but no G/G_i admits a morphism from 1. On the other hand, if G is finitely generated and

we are given a filtered diagram of G-sets A_i (indexed by a filtered category $\mathcal{I}$), and a morphism $1 \to \varinjlim_{\mathcal{I}} A_i$ (that is, an element a of this colimit whose stabilizer subgroup is the whole of G), then we can lift a to an element a_0 of some A_{i_0} (not necessarily fixed by the whole of G); now the stabilizer subgroups G_f of the elements $a_f = f(a_0) \in A_i$, as $f: i_0 \to i$ ranges over the co-slice category $i_0\backslash\mathcal{I}$, form a filtered diagram of subgroups of G whose union is the whole of G, so some G_f must be the whole of G. Thus we have shown that the canonical map $\varinjlim_{\mathcal{I}} \gamma_*(A_i) \to \gamma_*(\varinjlim_{\mathcal{I}} A_i)$ is surjective; but we already know it is injective, by compactness of $[G, \mathbf{Set}]$. A similar argument shows that *for a topological group G, the topos* $\mathbf{Cont}(G)$ *is strongly compact iff the top element G is inaccessible by directed joins in the poset of open subgroups of G.* In particular, if G is a compact topological group, then $\mathbf{Cont}(G)$ is strongly compact.

(c) Let $\mathcal{C}$ be a small cartesian category, and let T be a coverage on $\mathcal{C}$ whose covering families are all finite. Then the assertion that a given functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ is a T-sheaf is equivalent to saying that certain finite diagrams are limits in $\mathbf{Set}$; so using the fact that finite limits and filtered colimits commute in $\mathbf{Set}$, it is easy to see that $\mathbf{Sh}(\mathcal{C}, T)$ is closed in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ under filitered colimits, i.e. the inclusion functor preserves them. But the direct image functor $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}$ preserves all colimits, since it is simply given by evaluation at the terminal object of $\mathcal{C}$. So $\mathbf{Sh}(\mathcal{C}, T)$ is strongly compact. Thus *any coherent Grothendieck topos, in the sense of* D3.3.1, *is strongly compact.* In particular, any coherent locale (as defined in 2.4.3 above) is strongly compact.

For the moment we refrain from giving any further explicit examples of strongly compact locales; but, once we have developed the theory of locally compact locales in Section C4.1, we shall be able to show that any stably locally compact locale (in particular, any compact Hausdorff locale) is strongly compact (see 4.1.14). Indeed, it is true more generally that any compact Hausdorff topos is strongly compact; but we shall postpone the proof to 5.3.15, since it requires descent theory. (The converse is false, as may be seen by comparing 3.4.1(b) with 3.2.24, or by considering sheaves on a coherent locale which is not Hausdorff.)

Definition 3.4.2 Following I. Moerdijk and J. Vermeulen [858], we shall call a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ *tidy* if it makes $\mathcal{F}$ into a strongly compact $\mathcal{E}$-topos, i.e. if the direct image functor f_* preserves filtered $\mathcal{E}$-indexed colimits. (The term 'proper' has also been used for this notion by some authors; but since it is not a conservative extension of the well-established localic notion of propriety, we feel that the latter term is best used for the class of morphisms studied in Section C3.2.) Explicitly, this means that, given any object A of $\mathcal{E}$

and any filtered internal category $\mathbb{C}$ in $\mathcal{E}/A$, the square

$$\begin{tikzcd}
\mathcal{F}/f^*A \arrow[r, "f/A"] \arrow[d, "\infty"] & \mathcal{E}/A \arrow[d, "\infty"] \\
{[f^*\mathbb{C}, \mathcal{F}/f^*A]} \arrow[r, "{[\mathbb{C}, f/A]}"] & {[\mathbb{C}, \mathcal{E}/A]}
\end{tikzcd}$$

satisfies the Beck–Chevalley condition.

As in 3.2.12, we could have phrased the definition in terms of weakly filtered categories in $\mathcal{E}$, rather than filtered categories in slices of $\mathcal{E}$. It is again immediate from the definition that any stable right Beck–Chevalley morphism, in the sense of 2.4.16, is tidy; we shall prove the converse in 3.4.11 below.

We note in passing that Example 3.4.1(b) suffices to show that hyperconnected morphisms are not necessarily tidy. However, we have the following stability properties of tidy morphisms:

Lemma 3.4.3 *Given a commutative triangle of geometric morphisms*

$$\begin{tikzcd}
\mathcal{G} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\
 & \mathcal{E}
\end{tikzcd} ,$$

(i) *if f and g are tidy, then h is tidy;*

(ii) *if h is tidy and g is connected, then f is tidy;*

(iii) *if h is tidy and f is a closed inclusion, then g is tidy.*

Proof (i) and (ii) are just like the first two parts of 3.2.16, replacing the directed poset $\mathbb{I}$ by a filtered internal category $\mathbb{C}$ in $\mathcal{E}/A$, and in (ii) using the fact that the morphism $[h^*\mathbb{C}, \mathcal{G}/h^*A] \to [f^*\mathbb{C}, \mathcal{F}/f^*A]$ induced by g is connected (so that every object of its codomain is isomorphic to one in the image of its direct image functor). For (iii) we use the argument hinted at in the remark following 3.2.16: given a filtered internal category $\mathbb{D}$ in $\mathcal{F}$, the category $\mathbb{C} = f_*\mathbb{D}$ is filtered in $\mathcal{E}$ by A4.5.4 and the fact that regular functors preserve filtered

categories, and $\mathbb{D} \cong f^*\mathbb{C}$ since f is an inclusion. So in the diagram

$$\begin{tikzcd}
{[g^*\mathbb{D}, \mathcal{G}]} \arrow[r, "{[\mathbb{D}, g]_*}"] \arrow[d] & {[\mathbb{D}, \mathcal{F}]} \arrow[r, "{[\mathbb{C}, f]_*}"] \arrow[d] & {[\mathbb{C}, \mathcal{E}]} \arrow[d] \\
\mathcal{G} \arrow[r, "g_*"] & \mathcal{F} \arrow[r, "f_*"] & \mathcal{E}
\end{tikzcd}$$

where the vertical arrows are colimit functors, the outside and the right-hand square commute up to isomorphism (the latter since closed inclusions are tidy; for this, we may either use 3.4.1(c) and the fact that closed sublocales of the terminal locale are coherent, or 'internalize' the argument of A4.5.4 to show that the direct image f_* preserves arbitrary connected $\mathcal{E}$-indexed colimits, and in particular filtered ones), and f_* is conservative; so the Beck–Chevalley transformation in the left-hand square must be an isomorphism. The same argument applied in an arbitrary slice $\mathcal{F}/B$ of $\mathcal{F}$ (which is of course a closed subtopos of $\mathcal{E}/f_*B$) proves the result. □

Once we have proved the stability of tidy morphisms under pullback, we shall be able to delete the hypothesis that f is closed from 3.4.3(iii), by an argument like that in the proof of 3.2.16(iii). However, since tidy morphisms (are proper, and hence) have closed images, the particular case proved above suffices to show that a geometric morphism is tidy iff both parts of its surjection–inclusion factorization are tidy. It also follows from 3.4.3(ii) that the localic part of the hyperconnected–localic factorization of a tidy morphism is tidy. There seems to be no reason why the hyperconnected part should be tidy; but we shall see in 3.4.15 below that there is a canonical factorization of tidy morphisms into tidy factors which are respectively connected and localic, rather like the factorization of locally connected morphisms provided by 3.3.5(i).

The following criterion for strong compactness, due to K. R. Edwards [329], is often useful:

Lemma 3.4.4 *A* **Set**-*topos $\mathcal{E}$ is strongly compact iff it is compact and satisfies the following condition:*

(E) *Given any well-supported object A of $\mathcal{E}$, and a directed family $(R_i \mid i \in I)$ of equivalence relations on A whose union is the whole of $A \times A$, there exists $i \in I$ and a well-supported subobject $B \rightarrowtail A$ such that $B \times B \leq R_i$ as subobjects of $A \times A$.*

Proof For the necessity, suppose given A and R_i as in condition (E). Then the quotients A/R_i form a directed diagram in $\mathcal{E}$ whose colimit is 1, so if $\mathcal{E}(1, -)$ preserves filtered colimits then some A/R_i must admit a morphism from 1. Form

the pullback

$$\begin{tikzcd} B \arrow[r, tail] \arrow[d, two heads] & A \arrow[d, two heads] \\ 1 \arrow[r, tail] & A/R_i \end{tikzcd} ;$$

then the object B has the required properties.

Conversely, suppose the condition is satisfied, and let $(A_c \mid c \in \text{ob } \mathcal{C})$ be the vertices of a filtered diagram in $\mathcal{E}$. Since $\mathcal{E}$ is compact, we know that the comparison map $\varinjlim_{\mathcal{C}} \mathcal{E}(1, A_c) \to \mathcal{E}(1, \varinjlim_{\mathcal{C}} A_c)$ is injective, so we have only to prove that it is surjective. So suppose given $f: 1 \to \varinjlim_{\mathcal{C}} A_c$. Form the pullbacks

$$\begin{tikzcd} B_c \arrow[r] \arrow[d] & A_c \arrow[d] \\ 1 \arrow[r, "f"] & \varinjlim_{\mathcal{C}} A_c \end{tikzcd} ;$$

since colimits in $\mathcal{E}$ are stable under pullback, we have $1 \cong \varinjlim_{\mathcal{C}} B_c$. In particular, 1 is the union of the supports of the B_c, so by compactness some B_{c_0} must be well-supported. Now we restrict our attention to the (filtered) co-slice category $c_0 \backslash \mathcal{C}$; for each object $\alpha: c_0 \to c$ of this category, let R_α be the kernel-pair of $B_{c_0} \to B_c$. Then the R_α form a directed family of equivalence relations whose union is the whole of $B_{c_0} \times B_{c_0}$, so by assumption there exists some $B_{c_0} \to B_{c_1}$ whose kernel-pair contains the whole of $B' \times B'$, for some well-supported subobject $B' \rightarrowtail B_{c_0}$. But this is equivalent to saying that the composite $B' \rightarrowtail B_{c_0} \to B_{c_1}$ factors through 1; now the composite $1 \to B_{c_1} \to A_{c_1}$ is an element of $\mathcal{E}(1, A_{c_1})$ mapping onto the given element f. □

The proof of 3.4.4 is constructive, and so can be internalized to an arbitrary topos $\mathcal{S}$ to give a criterion for tidiness of morphisms $\mathcal{E} \to \mathcal{S}$.

We now turn to the problem of obtaining a characterization of bounded tidy morphisms in terms of sites. It appears that here, in contrast to the situation for proper maps (cf. the remarks after 3.2.19), there is no way of avoiding the assumption that we are working over a topos with a natural number object; for we need to be able to construct the equivalence relation generated by a subobject of a square (i.e. an object of the form $U \times U$ of our site) as a countable union. It therefore seems sensible to restrict to sites whose underlying categories are pretoposes, rather than merely coherent categories as in Section C3.2; this assumption (or at least the assumption that $\mathcal{C}$ is positive as a coherent category) has an appreciable simplifying effect on the condition we obtain.

Our basic strategy is to rewrite the condition (E) of 3.4.4 in terms of our site $(\mathcal{C}, J)$. Of course, in addition to assuming that $\mathcal{C}$ is a pretopos, we shall

assume as in Section C3.2 that J is the join of the coherent coverage P and a dm-coverage D, such that the terminal object of $\mathcal{C}$ is D-irreducible. Given a subobject $V \rightarrowtail U \times U$ in $\mathcal{C}$, we shall write $V^{(1)}$ for the reflexive and symmetric closure of V (that is, the union of V, its opposite and the diagonal) and $V^{(n)}$ $(n > 1)$ for the nth iterate (in $\mathbf{Rel}(\mathcal{C})$) of $V^{(1)}$, and refer to the sieve on $U \times U$ generated by the $V^{(n)} \rightarrowtail U \times U$ as the *equivalence sieve* associated with V; it is of course a dm-sieve.

Proposition 3.4.5 *A Grothendieck topos $\mathcal{E}$ is strongly compact iff there exists a site of definition $(\mathcal{C}, J)$ for $\mathcal{E}$ such that $\mathcal{C}$ is a (small) pretopos, $J = P \vee D$ is a coverage satisfying the hypotheses of* 3.2.19(iv) *and, if U is any well-supported object of $\mathcal{C}$ and R any D-covering sieve on $U \times U$, there exists $a: U' \rightarrowtail U$ such that U' is well-supported, and a morphism $e: V \rightarrowtail U \times U$ in R such that the pullback along $a \times a: U' \times U' \rightarrowtail U \times U$ of the equivalence sieve associated with e is D-covering.*

Proof First suppose $\mathcal{E} = \mathbf{Sh}(\mathcal{C}, J)$ where $\mathcal{C}$ satisfies the given conditions. By 3.2.19 we know $\mathcal{E}$ is compact, so we must show that it satisfies condition (E). So suppose given a well-supported object A of $\mathcal{E}$, and a directed family of equivalence relations $(R_i \mid i \in I)$ as in the statement of (E). We note first that there is a well-supported object U of $\mathcal{C}$ such that $l(U)$ admits a morphism to A (equivalently, $A(U)$ is inhabited); for the objects U such that $A(U)$ is inhabited must cover the terminal object, and by compactness there is a finite subfamily of them which cover, and we can form the coproduct of this family in our site. Now $l(U \times U) \cong l(U) \times l(U)$ is covered in $\mathcal{E}$ by the subobjects which are the pullbacks of the R_i, so we can find a D-covering sieve on $U \times U$ whose image under l refines this cover. Let $a: U' \rightarrowtail U$ and $e: V \rightarrowtail U \times U$ be as in the statement of the proposition. By construction, $l(V) \rightarrowtail l(U \times U) \to A \times A$ factors through some $R_i \rightarrowtail A \times A$, and since the latter is an equivalence relation it follows that $l(V^{(n)}) \rightarrowtail l(U \times U) \to A \times A$ also factors through R_i, for all n. But the pullbacks of the $l(V^{(n)}) \rightarrowtail l(U \times U)$ along $l(U' \times U') \rightarrowtail l(U \times U)$ are jointly epic; so if we take $B \rightarrowtail A$ to be the image of $l(U') \rightarrowtail l(U) \to A$, then $B \times B$ is contained in R_i (and $B \to 1$ is epic, since $l(U') \to 1$ is epic). So condition (E) is verified.

Conversely, suppose $\mathcal{E}$ is strongly compact; let $\mathcal{C}$ be a small full generating subcategory of $\mathcal{E}$ which is closed under finite limits and colimits, and J the coverage induced on it. By 3.2.18, we can decompose J as $P \vee D$ in the required manner. To verify the condition in the statement of the Proposition, suppose given U and R as stated there. For each $e: V \rightarrowtail U \times U$ in R, let $\overline{V} \rightarrowtail l(U) \times l(U)$ be the equivalence relation (in $\mathcal{E}$) generated by $l(V) \rightarrowtail l(U) \times l(U)$: of course, it is the union in $\mathrm{Sub}_{\mathcal{E}}(l(U) \times l(U))$ of the $l(V^{(n)}) \rightarrowtail l(U) \times l(U)$. By condition (E), there exists $B \rightarrowtail l(U)$ such that $B \to 1$ is epic and $B \times B$ is contained in $\overline{V}$ for some V. As in the first part of the proof, we may now choose $U' \rightarrowtail U$ in $\mathcal{C}$ such that U' is well-supported and $l(U') \rightarrowtail l(U)$ factors through $B \rightarrowtail l(U)$; so $l(U') \times l(U')$ is also contained in $\overline{V}$, whence $U' \times U'$ must be covered by the

directed family of monomorphisms $(V^{(n)} \cap (U' \times U') \rightarrowtail U' \times U' \mid n \geq 1)$. So the latter generates the required D-covering sieve. $\square$

We shall call a monomorphism $V \rightarrowtail U \times U$ in $\mathcal{C}$, where U is well-supported, an *effective subobject* (relative to a given coverage on $\mathcal{C}$) if the equivalence sieve associated with it covers $U' \times U'$ for some well-supported $U' \rightarrowtail U$. Thus the condition in the statement of 3.4.5 says that every dm-cover of a well-supported square contains an effective subobject. (Warning: this notion of effectivity is not simply a property of the monomorphism in question; it also depends on the particular decomposition of its codomain as a square $U \times U$ – if it happens that $U \times U$ coincides with $U' \times U'$ in $\mathcal{C}$, the effective subobjects of the two squares may be entirely different. Thus the set E appearing in the next lemma should really be viewed as a set of pairs (U, e) where U is a well-supported object and $e \colon V \rightarrowtail U \times U$; however, to simplify the notation we shall treat it as if its members were simply the monomorphisms e.)

Lemma 3.4.6 *Let $\mathcal{C}$ be a pretopos; let D be a (sifted, but not necessarily Grothendieck) dm-coverage on $\mathcal{C}$, and suppose we have a family E of distinguished monomorphisms into well-supported squares with the following properties:*

(i) *Every $(V \rightarrowtail U \times U)$ in E is effective (with respect to the coverage D).*

(ii) *If $(V \rightarrowtail U \times U) \in E$ and $R \in D(V)$, then there exists $(V' \rightarrowtail V) \in R$ such that the composite $V' \rightarrowtail V \rightarrowtail U \times U$ is still in E.*

(iii) *For every well-supported object U, the identity morphism $U \times U \rightarrowtail U \times U$ belongs to E.*

Then the Grothendieck coverage $\tilde{D}$ generated by D has the property that every $\tilde{D}$-covering sieve on a well-supported square contains a monomorphism in E. In particular, the topos $\mathbf{Sh}(\mathcal{C}, P \vee \tilde{D})$ (where P is the coherent coverage) is strongly compact.

Proof Consider the set T of all dm-sieves R on objects of $\mathcal{C}$ with the property that, given a morphism $V \to b(R)$ and a monomorphism $V \rightarrowtail U \times U \in E$, there exists $V' \rightarrowtail V$ such that $V' \rightarrowtail V \rightarrowtail U \times U$ is still in E and the composite $V' \rightarrowtail V \to b(R)$ belongs to R. We must show that this set is a Grothendieck coverage; it clearly contains D, by condition (ii), so it will then follow that it contains $\tilde{D}$.

The form of the definition of T automatically ensures that it satisfies the pullback-stability condition (C); and condition (M) is guaranteed by condition (iii) in the statement of the lemma. So we have to verify condition (L). Let R, S be dm-sieves on W such that $R \in T(W)$ and $a^*(S) \in T(\text{dom } a)$ for all $a \in R$, and suppose given $a \colon V \to W$ and $e \colon V \rightarrowtail U \times U \in E$. By assumption, we can find $V' \rightarrowtail V$ such that $V' \rightarrowtail V \to W$ is in R and $V' \rightarrowtail U \times U$ is still in E. Now the pullback of S along the composite $a' \colon V' \rightarrowtail V \to W$ is in $T(V')$, so applying the definition of T to it we obtain $V'' \rightarrowtail V' \in a'^*(S)$ such that $V'' \rightarrowtail U \times U$ is

still in E. Thus the composite $V'' \rightarrowtail V \to W$ is in S, and we have verified that $S \in T(W)$.

The second assertion of the lemma now follows from condition (i) and 3.4.5, since if the morphisms in E are D-effective they are certainly $\tilde{D}$-effective. □

Theorem 3.4.7 *Suppose given a pullback square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}$, *where* f *is bounded and tidy, and* $\mathcal{E}$ *has a natural number object. Then* k *is tidy, and the Beck–Chevalley natural transformation* $g^* f_* \to k_* h^*$ *is an isomorphism.*

Proof Let $(\mathbb{C}, J)$ be an internal site for $\mathcal{F}$ over $\mathcal{E}$ satisfying the hypotheses of 3.4.5. Then, as in the proof of 3.2.21(i), we may obtain a site for $\mathcal{H}$ over $\mathcal{G}$ by equipping the pretopos $g^*\mathbb{C}$ with the coverage which is the join of its coherent coverage and the Grothendieck closure $\widetilde{g^\# D}$ of $g^\# D$. Let $E \rightarrowtail C_0 \times C_1$ be the set of pairs $(U, e \colon V \rightarrowtail U \times U)$ such that U is well-supported and e is a D-effective subobject; then clearly E satisfies the conditions of 3.4.6. We claim that $g^* E$ satisfies the same conditions with respect to the coverage $g^\# D$ on $g^*\mathbb{C}$; the fact that it satisfies the second and third conditions is obvious, and the first condition follows from the fact that, since g^* preserves the natural number object, it also preserves the construction of the equivalence sieve associated with a monomorphism into a square. Hence $(g^*\mathbb{C}, g^\# P \vee \widetilde{g^\# D})$ satisfies the hypotheses of 3.4.5, and so k is tidy.

For the Beck–Chevalley condition, consider an object of $\mathcal{F}$, which we may think of as a discrete fibration $\mathbb{F} \to \mathbb{C}$ satisfying the sheaf axiom for all covers in J. Then $h^*\mathbb{F}$ is simply the reflection of $g^*\mathbb{F} \to g^*\mathbb{C}$ in the subcategory of sheaves on $g^*\mathbb{C}$, and the Beck–Chevalley transformation $g^* f_*(\mathbb{F}) \to k_* h^*(\mathbb{F})$ is simply the component of the unit map at the terminal object 1 of $g^*\mathbb{C}$. By 3.2.21, we already know this morphism is monic, so we need only show that it is epic.

First we note that $\mathbb{F}$ satisfies the condition which says that, given a well-supported object U of $\mathbb{C}$, an element $x \in F(U)$ and a morphism $e = (e_1, e_2) \colon V \rightarrowtail U \times U$ in E such that $F(b_1)(x) = F(b_2)(x)$, there exists a well-supported $U' \rightarrowtail U$ and an element $y \in F(1)$ such that $F(U' \rightarrowtail U)(x) = F(U' \to 1)(y)$ in $F(U')$. For, since $\mathbb{F}$ is a sheaf for the coherent coverage, the fact that the elements $F(\pi_1)(x)$ and $F(\pi_2)(x)$ of $F(U \times U)$ become equal on restriction to V implies that they become equal on restriction to $V^{(n)}$ for each n; but these subobjects generate a sieve whose pullback to $U' \times U'$ is D-covering, and hence $F(U' \rightarrowtail U)(x)$ has equal images under the two projections $U' \times U' \rightrightarrows U'$. But this says that the restriction of x to $F(U')$ defines a compatible family for the

P-covering sieve generated by $U' \twoheadrightarrow 1$; so it is the image of a unique $y \in F(1)$. Now the condition at the beginning of this paragraph may be expressed by a coherent sequent (cf. D1.1.5) in the internal language of $\mathcal{E}$; hence $g^*\mathbb{F}$ satisfies the same condition in $\mathcal{G}$.

Now let $\overline{\mathbb{F}} \to g^*\mathbb{C}$ denote the associated sheaf of $g^*\mathbb{F}$, and suppose given an element $x \in \overline{F}(1)$. By the definition of the associated sheaf given in the proof of 2.2.6, x is determined by a family of elements $(x_i \in g^*(F)(U_i) \mid i \in I)$, where $(U_i \to 1 \mid i \in I)$ is a covering family, such that the x_i are 'locally compatible', i.e., for each pair (i, j), the elements $g^*F(\pi_1)(x_i)$ and $g^*F(\pi_2)(x_j)$ agree on a cover of $U_i \times U_j$. Now, by compactness of $g^*\mathbb{C}$ and the fact that it has finite coproducts, we may replace the family $(x_i \mid i \in I)$ by a single element $y \in g^*F(U)$ for some $U \twoheadrightarrow 1$ in $g^*\mathbb{C}$; moreover, it is easy to see that the family S of all morphisms $(e_1, e_2)\colon V \to U \times U$ for which $g^*F(e_1)(y) = g^*F(e_2)(y)$ is a dm-sieve on $U \times U$, since $g^*\mathbb{F}$ is a sheaf for the coherent coverage. Hence this covering sieve belongs to $g^{\#}D$. So, by the proof of 3.4.6, we can find a monomorphism $e\colon V \rightarrowtail U \times U$ in S which belongs to $g^*(E)$; but, by the argument of the previous paragraph, this means that we can find a well-supported $U' \rightarrowtail U$ such that the restriction y' of y to $g^*(F)(U')$ derives from an element $z \in g^*(F)(1)$. Now the image of z in $\overline{F}(1)$ must equal x, since they agree on the covering $U' \twoheadrightarrow 1$ and $\overline{\mathbb{F}}$ is a sheaf; so we have verified that $g^*F(1) \to \overline{F}(1)$ is epic. $\square$

As usual, we have a second proof of the pullback-stability theorem for tidy morphisms, in which we assume boundedness for the morphism along which we are pulling back, and factor it as an inclusion followed by one of the form $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$. However, this proof differs from those in the last three sections, in that we have to assume boundedness for the tidy morphism as well; this is because we shall need to appeal to Theorem 3.3.14, which we have proved only for bounded morphisms. Nevertheless, the second proof seems worth giving, since it does not require the assumption that the toposes involved have natural number objects, and it is therefore applicable in contexts where 3.4.7 is not.

In fact, the first step in the proof does not require the tidy morphism to be bounded.

Lemma 3.4.8 *Let $f : \mathcal{F} \to \mathcal{E}$ be a tidy geometric morphism, and let $\mathbb{C}$ be an internal category in $\mathcal{E}$. Then the induced morphism $[\mathbb{C}, f]\colon [f^*\mathbb{C}, \mathcal{F}] \to [\mathbb{C}, \mathcal{E}]$ is tidy, and the pullback square*

$$\begin{tikzcd}
{[f^*\mathbb{C}, \mathcal{F}]} \arrow[r, "{[\mathbb{C}, f]}"] \arrow[d] & {[\mathbb{C}, \mathcal{E}]} \arrow[d] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

satisfies the Beck–Chevalley condition.

Proof The proof that $[\mathbb{C}, f]$ is tidy is exactly like the corresponding part of the second proof of 3.2.21: given a filtered internal category $\mathbb{I}$ in $[\mathbb{C}, \mathcal{E}]$, we let $\mathbb{J}$ denote its underlying (filtered) category in $\mathcal{E}/C_0$, and form the cube displayed in the proof of 3.2.21. As before, the fact that the top face commutes up to isomorphism follows from the fact that the bottom face does so and the vertical edges reflect isomorphisms. The Beck–Chevalley condition is easily verified by direct calculation, without using the tidiness of f (which is not surprising, since $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$ is locally connected). □

Lemma 3.4.9 *Let $f: \mathcal{F} \to \mathcal{E}$ be a bounded tidy morphism; let j be a local operator on $\mathcal{E}$, and let j' be the local operator on $\mathcal{F}$ for which the square*

$$\begin{tikzcd} \mathbf{sh}_{j'}(\mathcal{F}) \arrow[r, "k"] \arrow[d, "h"] & \mathbf{sh}_j(\mathcal{E}) \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

is a pullback. Then the Beck–Chevalley natural transformation $g^* f_* \to k_* h^*$ is an isomorphism.

Proof By A4.3.12, k_* is simply the restriction of f_* to sheaves; so the lemma is equivalent to the statement that f_* commutes up to isomorphism with the associated sheaf functors for j and j'. But, as we remarked at the end of Section A4.4, the associated j-sheaf of an object A may be constructed as A^{++}, where A^+ is the colimit of an internal diagram $\hat{A}$ of shape $\mathbf{J}^{\mathrm{op}}$, and $\mathbf{J} = (J_1 \rightrightarrows J)$ is the sub-poset of $\mathbf{\Omega}$ determined by the object J (this is the internalization of the construction of this functor which we gave in the proof of 2.2.6; we shall not give the details here, but the interested reader may find them in [501] or [504]). So it suffices to show that the diagram

$$\begin{tikzcd} \mathcal{F} \arrow[r, "f_*"] \arrow[d, "+"] & \mathcal{E} \arrow[d, "+"] \\ \mathcal{F} \arrow[r, "f_*"] & \mathcal{E} \end{tikzcd}$$

commutes up to isomorphism. But the first and third conditions in the definition of a local operator ensure that $\mathbf{J}$ is an internal meet-semilattice in $\mathcal{E}$, and so $\mathbf{J}^{\mathrm{op}}$ is filtered; hence this is almost immediate from the definition of a tidy morphism. The only reason why it is not quite immediate is that, to compute the functor $+$ in $\mathcal{F}$, we have to form a colimit over the internal category $\mathbf{J}'^{\mathrm{op}}$, rather than

$f^*\mathbf{J}^{\mathrm{op}}$. However, if B is an arbitrary object of $\mathcal{F}$ and we form the appropriate diagram $\hat{B}$ of shape $\mathbf{J}'^{\mathrm{op}}$, it is easily verified from the description of $\hat{A}$ in the proof of A4.4.10 that

$$\begin{tikzcd} \widehat{f^*B} \arrow[r] \arrow[d] & f_*(\hat{B}) \arrow[d] \\ J \arrow[r] & f_*J' \end{tikzcd}$$

is a pullback, where the bottom edge of the square is the transpose of the canonical morphism $f^*J \to J'$ in $\mathcal{F}$. So, to complete the proof, it suffices by B2.5.12 to show that $f^*\mathbf{J} \to \mathbf{J}'$ is an initial functor. But that is precisely what 3.3.14 tells us. $\square$

Theorem 3.4.10 *Suppose given a pullback square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}$, *where* f *is bounded and tidy, and* g *is bounded. Then* k *is tidy, and the Beck–Chevalley natural transformation* $g^*f_* \to k_*h^*$ *is an isomorphism.*

Proof By 3.4.8, we may reduce to the case when g is an inclusion; and the Beck–Chevalley condition for this case is proved in 3.4.9. So it remains to prove that k is tidy, which we shall do by showing that it is a stable right Beck–Chevalley morphism.

Suppose given a bounded morphism $r\colon \mathcal{G}' \to \mathcal{G}$. Then the composite gr is also bounded, so we may factor it as

$$\mathcal{G}' \xrightarrow{g'} \mathcal{E}' \xrightarrow{p} \mathcal{E}$$

where g' is an inclusion and $\mathcal{E}' \simeq [\mathbb{C}, \mathcal{E}]$ for some internal category $\mathbb{C}$ in $\mathcal{E}$. Now

we may form the cube

$$
\begin{tikzcd}
\mathcal{H}' \arrow[rr, "k'"] \arrow[dr, "s"] \arrow[dd, "h'"'] & & \mathcal{G}' \arrow[dr, "r"] \arrow[dd, "g'" near start] & \\
& \mathcal{H} \arrow[rr, "k" near start] \arrow[dd, "h" near end] & & \mathcal{G} \arrow[dd, "g"] \\
\mathcal{F}' \arrow[rr, "f'" near end] \arrow[dr, "q"'] & & \mathcal{E}' \arrow[dr, "p"] & \\
& \mathcal{F} \arrow[rr, "f"] & & \mathcal{E}
\end{tikzcd}
$$

in which all faces with horizontal edges are pullbacks, and the other two faces commute. Moreover, the Beck–Chevalley condition holds for the bottom face by 3.4.8, and for the front and back faces by 3.4.9, so an easy diagram-chase shows that the canonical natural transformation $r^*k_*h^* \to k'_*s^*h^*$ is an isomorphism. But h is an inclusion, so h^* is essentially surjective on objects; hence $r^*k_* \to k'_*s^*$ is an isomorphism.

Thus we have shown that k is a RBC morphism in the sense of 2.4.16. But the same diagram shows that any pullback of k along a bounded morphism is expressible as a pullback of a tidy morphism along an inclusion; so it too is a RBC morphism. Hence k is a SRBC morphism; as we remarked after Definition 3.4.2, this trivially implies that it is tidy. □

Corollary 3.4.11 *A bounded geometric morphism $f: \mathcal{F} \to \mathcal{E}$ is tidy iff it is a stable right Beck–Chevalley morphism in the sense of* 2.4.16.

Proof One direction is immediate from the definition; the converse follows from 3.4.10 (or from 3.4.7, if $\mathcal{E}$ has a natural number object). □

We do not yet have a 'dual' for the part of 3.3.15 which says that connected locally connected morphisms are stable under pullback. We could obtain this by observing that connectedness is always preserved under pullback along morphisms of the form $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$, and then using the Beck–Chevalley condition to show that connectedness of tidy morphisms is preserved under pullback along inclusions. However, we shall take an alternative route, which involves discussing tidiness in the context of the *pure–entire factorization* of a geometric morphism. We have not met this factorization before (though we met entire morphisms

in B4.4.21(a)), but it is easily described: recall that a morphism $f: \mathcal{F} \to \mathcal{E}$ is said to be entire if it is localic and the corresponding internal locale X in $\mathcal{E}$ is compact and zero-dimensional (i.e. $\mathcal{O}(X)$ is generated by its complemented elements). It is well known that compact zero-dimensional locales are coherent; in fact they are precisely those whose frames are of the form IB for a Boolean algebra B (cf. [520, II 4.2]). It thus follows immediately from 2.4.3 that entire morphisms are stable under pullback, and from 3.4.1(c) that they are tidy.

We recall also that in any topos we have $\Omega \cong I(2)$, where $2 = 1 \amalg 1$ is the Boolean algebra of complemented elements of Ω. We define a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ to be *pure* if f_* preserves 2; more explicitly, if the canonical morphism $2_{\mathcal{E}} \to f_*(2_{\mathcal{F}})$ (the unit of $(f^* \dashv f_*)$ at the object 2) is an isomorphism.

Lemma 3.4.12

(i) *A morphism f is pure iff f_* preserves finite coproducts.*

(ii) *Hyperconnected morphisms are pure.*

Proof (i) One direction is immediate since $2 = 1 \amalg 1$ and f_* preserves 1. For the converse, we note first that a pure morphism f must be dense (i.e. f_* must preserve 0), since f_* preserves the pullback

$$\begin{tikzcd} 0 \arrow[r] \arrow[d] & 1 \arrow[d, "\nu_2"] \\ 1 \arrow[r, "\nu_1"] & 2 \end{tikzcd} .$$

For binary coproducts, we similarly use the fact that f_* preserves the pullbacks.

$$\begin{tikzcd} A \arrow[r, "\nu_1"] \arrow[d] & A \amalg B \arrow[d] & B \arrow[l, "\nu_2"'] \arrow[d] \\ 1 \arrow[r, "\nu_1"] & 2 & 1 \arrow[l, "\nu_2"'] \end{tikzcd} .$$

(ii) If $f: \mathcal{F} \to \mathcal{E}$ is hyperconnected, then for any object A of $\mathcal{E}$ we have an isomorphism $\mathrm{Sub}_{\mathcal{E}}(A) \cong \mathrm{Sub}_{\mathcal{F}}(f^*A)$, which clearly restricts to an isomorphism between their sublattices of complemented elements. So morphisms $A \to 2_{\mathcal{E}}$ correspond bijectively to morphisms $A \to f_*(2_{\mathcal{F}})$. $\square$

Proposition 3.4.13

(i) *Pure morphisms are orthogonal to entire morphisms: that is, given a commutative square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

where h is pure and g is entire, there exists (uniquely up to canonical isomorphism) a morphism $l: \mathcal{F} \to \mathcal{G}$ making the two triangles commute.

(ii) *Every geometric morphism can be factored as a pure morphism followed by an entire morphism, and the factorization is unique up to canonical equivalence.*

Proof (i) Since entire morphisms are stable under pullback, it suffices to show that if $g: \mathcal{G} \to \mathcal{E}$ is entire and $h: \mathcal{H} \to \mathcal{E}$ is pure, then any morphism $k: \mathcal{H} \to \mathcal{G}$ over $\mathcal{E}$ factors through a unique section of g. But this is easy, since k corresponds to a frame homomorphism $g_*(\Omega_{\mathcal{G}}) \to h_*(\Omega_{\mathcal{H}})$ in $\mathcal{E}$, and hence to a lattice homomorphism $g_*(2_{\mathcal{G}}) \to h_*(\Omega_{\mathcal{H}})$; and the latter must factor (uniquely) through the Boolean algebra $h_*(2_{\mathcal{H}}) \cong 2_{\mathcal{E}}$ of complemented elements of $h_*(\Omega_{\mathcal{H}})$.

(ii) Given an arbitrary morphism $f: \mathcal{F} \to \mathcal{E}$, let X be the locale in $\mathcal{E}$ defined by $\mathcal{O}(X) = I(f_*(2_{\mathcal{F}}))$, and let $\mathcal{G} = \mathbf{Sh}_{\mathcal{E}}(X)$. Then $\mathcal{G}$ is clearly entire over $\mathcal{E}$; and the canonical morphism $f_*(2_{\mathcal{F}}) \rightarrowtail f_*(\Omega_{\mathcal{F}})$ induces a frame homomorphism $I(f_*(2_{\mathcal{F}})) \to f_*(\Omega_{\mathcal{F}})$ (cf. 1.1.3) and hence a geometric morphism $h: \mathcal{F} \to \mathcal{G}$ over $\mathcal{E}$. We claim that h is pure.

To prove this, we may use 3.4.12(ii) to reduce to the case when f is localic, i.e. $\mathcal{F} \simeq \mathbf{Sh}_{\mathcal{E}}(X)$ where X is the internal locale defined by $\mathcal{O}(X) = f_*(\Omega_{\mathcal{F}})$. In terms of this description, $\Omega_{\mathcal{F}}$ is the sheaf on $\mathcal{O}(X)$ corresponding to the discrete fibration $\mathcal{O}(X)_1 \to \mathcal{O}(X)$ whose domain is the order-relation on $\mathcal{O}(X)$, and the mapping sends a pair (U, V) of opens in X with $U \leq V$ to V (cf. A2.1.8). Clearly, the object $2_{\mathcal{F}}$ corresponds to the subobject

$$\{(U, V) \mid (\exists W : \mathcal{O}(X))((U \cap W = \emptyset) \text{ and } (U \cup W = V))\}$$

of $\mathcal{O}(X)_1$. Now the direct image of the geometric morphism h is simply given by pullback along the frame homomorphism $I(f_*(2_{\mathcal{F}})) \to \mathcal{O}(X)$, which of course sends an ideal of complemented elements of $\mathcal{O}(X)$ to its join. If we use the alternative representation of objects of $\mathcal{G}$ as sheaves for the coherent coverage on $f_*(2_{\mathcal{F}})$ itself, we see that $h_*(2_{\mathcal{F}})$ corresponds to the functor whose value at a complemented open U is the set of complemented opens below U. But this is precisely the object $2_{\mathcal{G}}$; so h is pure.

This establishes the existence of the factorization; the uniqueness follows from (i). $\square$

Note that it follows from 3.4.13 that a composite of entire morphisms is entire, as we claimed in B4.4.22. (In fact it would not be hard to prove this directly from the definition.)

Generalizing 3.4.12(ii), connected morphisms are pure; in fact

Lemma 3.4.14 *For a geometric morphism $f: \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is connected.*

(ii) *f_* preserves $\mathcal{E}$-indexed coproducts.*

(iii) *The pullback of f along any entire morphism with codomain $\mathcal{E}$ is pure.*

Proof (ii) $\Rightarrow$ (i) is immediate, since the assertion that f_* preserves $\mathcal{E}$-indexed coproducts in particular says that, for any object A of $\mathcal{E}$, the canonical morphism $\Sigma_A(f/A)_*(1_{f^*A}) \to f_*\Sigma_{f^*A}(1_{f^*A})$ is an isomorphism; but this canonical morphism is just the unit map $A \to f_*f^*A$.

(i) $\Rightarrow$ (ii): Conversely, if f is connected, then for any $h: A \to B$ in $\mathcal{E}$ the two ways round the diagram

$$\begin{tikzcd} \mathcal{F}/f^*A \arrow[r, "\Sigma_{f^*h}"] \arrow[d, "(f/A)_*"] & \mathcal{F}/f^*B \arrow[d, "(f/b)_*"] \\ \mathcal{E}/A \arrow[r, "\Sigma_h"] & \mathcal{E}/B \end{tikzcd}$$

both send an arbitrary $k: C \to f^*A$ to the composite $h \cdot (\eta_A)^{-1} \cdot (f_*k)$, where η is the unit of $(f^* \dashv f_*)$.

(i) $\Leftrightarrow$ (iii): Let $g: \mathcal{G} \to \mathcal{E}$ be an entire morphism, corresponding to an internal Boolean algebra B in $\mathcal{E}$, and form the pullback

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd} \quad .$$

Then the assertion that k is pure is equivalent to saying that the top and right edges of the above square are the pure–entire factorization of the diagonal, i.e. that $f_*h_*(2_{\mathcal{H}}) \cong B$ as Boolean algebras in $\mathcal{E}$. But h is also entire; indeed, by the proof of 2.4.3 we know that $h_*(2_{\mathcal{H}})$ is the Boolean algebra $f^*(B)$. Hence (iii) is equivalent to the assertion that the unit map $B \to f_*f^*(B)$ is an isomorphism for every object B of $\mathcal{E}$ which is the underlying object of a Boolean

algebra. But by A4.5.23 there exists a surjection $p\colon \mathcal{B} \to \mathcal{E}$ where $\mathcal{B}$ is a Boolean topos; so for any object A of $\mathcal{E}$ we have a monomorphism

$$A \rightarrowtail p_*p^*(A) \xrightarrow{p_*(\{\})} p_*(2^{p^*(A)})$$

from A to the underlying object of a Boolean algebra, and by applying the same construction to the cokernel-pair of this morphism we may represent A as the equalizer of a pair of maps between the underlying objects of Boolean algebras. Since f^* and f_* both preserve equalizers, the assertion above is thus equivalent to saying that $A \to f_*f^*(A)$ is an isomorphism for all A. □

The pure–entire factorization of an arbitrary morphism is not stable under arbitrary pullback; but if a pullback square

$$\begin{tikzcd}
\mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

satisfies the Beck–Chevalley condition $g^*f_* \cong k_*h^*$, then in particular we have $g^*f_*(2_{\mathcal{F}}) \cong k_*h^*(2_{\mathcal{F}}) \cong k_*(2_{\mathcal{H}})$, and so the pure–entire factorization of f pulls back to that of k. Hence we may conclude

Corollary 3.4.15

(i) *The pure–entire factorization of an arbitrary morphism is stable under pullback along locally connected morphisms.*

(ii) *The pure–entire factorization of a bounded tidy morphism is stable under arbitrary bounded pullback. Moreover, both halves of the factorization are tidy.*

(iii) *A bounded tidy morphism is pure iff it is connected.*

(iv) *Connected bounded tidy morphisms are stable under pullback.*

Proof (i) and the first assertion of (ii) are immediate from 3.3.15 and 3.4.10 respectively. For the second assertion of (ii), we recall that entire morphisms are tidy in any case; for the pure part, let

$$\mathcal{F} \xrightarrow{h} \mathcal{G} \xrightarrow{g} \mathcal{E}$$

be the pure–entire factorization of a tidy morphism $f: \mathcal{F} \to \mathcal{E}$. Then we have a diagram

$$\begin{tikzcd}
 & \mathcal{F} \arrow[r, "h"] \arrow[d, "{(1,h)}"'] & \mathcal{G} \arrow[d, "\Delta_g"] \\
\mathcal{F} \arrow[d, "f"] & \mathcal{F} \times_{\mathcal{E}} \mathcal{G} \arrow[l, "\pi_1"'] \arrow[r, "h \times 1"] \arrow[d, "\pi_2"] & \mathcal{G} \times_{\mathcal{E}} \mathcal{G} \\
\mathcal{E} & \mathcal{G} \arrow[l, "g"'] &
\end{tikzcd}$$

where the squares are pullbacks and the vertical composite in the middle is h; but f is tidy by assumption and Δ_g is tidy because it is a closed inclusion. So the result follows from the stability of tidy morphisms under composition and pullback.

(iii) now follows from (ii) and 3.4.14(iii), and (iv) is immediate from (ii) and (iii). □

Remark 3.4.16 An alternative proof of 3.4.15(iii), without the assumption of boundedness, may be given using 3.4.12(i) and 3.4.14(ii), plus the fact that arbitrary coproducts may be expressed as filtered colimits of finite ones (cf. D5.2.14).

It is also possible to give an alternative, more direct (but rather more long-winded!), proof of the fact that the pure part of the pure–entire factorization of a tidy morphism is tidy. The argument relies on the following result, which is of interest in its own right – it extends to entire morphisms something which we already knew for closed inclusions, by A4.5.4.

Lemma 3.4.17 *Let $f: \mathcal{F} \to \mathcal{E}$ be an entire morphism. Then f_* preserves coequalizers of reflexive pairs; in particular, it is a regular functor.*

Proof Suppose $\mathcal{F} \simeq \mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P)$, where $\mathbb{B}$ is an internal Boolean algebra in $\mathcal{E}$ and P is its coherent coverage. We claim that the inclusion $\mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P) \to [\mathbb{B}^{\mathrm{op}}, \mathcal{E}]$ preserves reflexive coequalizers; since the direct image functor $[\mathbb{B}^{\mathrm{op}}, \mathcal{E}] \to \mathcal{E}$ is given by evaluation at the top element of $\mathbb{B}$ (and hence preserves all colimits), this will suffice. But the coverage P is generated by the empty cover of the bottom element 0 and all two-element covers $(a \to a \vee b\,,\, b \to a \vee b)$ where $a \wedge b = 0$ (since any two-element cover may be refined to a disjoint one); hence we see that a diagram F on $\mathbb{B}^{\mathrm{op}}$ is a P-sheaf iff $F(0) \cong 1$ and $F(a \vee b) \cong F(a) \times F(b)$ whenever a and b are disjoint. Using the fact that binary products commute with reflexive coequalizers in a cartesian closed category (A1.2.12), it follows easily that if $F \rightrightarrows G$ is a reflexive pair of morphisms in $\mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P)$ then their

'pointwise' coequalizer (i.e. their coequalizer in $[\mathbb{B}^{\mathrm{op}}, \mathcal{E}]$) is a sheaf, and hence is their coequalizer in $\mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P)$. The second assertion of the lemma follows from the fact that any epimorphism in a topos may be presented as the coequalizer of its kernel-pair (A1.3.4). □

A further property of entire morphisms which we require is that, if $f: \mathcal{F} \to \mathcal{E}$ is entire and A is any well-supported object of $\mathcal{F}$, then the counit map $f^* f_* A \to A$ is epic. For if we represent $\mathcal{F}$ as $\mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P)$ where $\mathbb{B} = (B_1 \rightrightarrows B)$ is a Boolean algebra and P is its coherent coverage, then the assertion that $A \to 1$ is epic implies by the result just proved that $A(1)$ is inhabited (where 1 denotes the top element of B). Hence we can represent any $x \in A(b)$ as the restriction to b of an element of $A(1) = f_*(A)$, by choosing an arbitrary $y \in A(1)$ and then patching together x and the restriction of y to $A(\neg b)$; this defines an element of $f^* f_*(A)(1)$ whose restriction to $f^* f_*(A)(b)$ maps onto x. Thus $f^* f_*(A) \to A$ is already epic in $[\mathbb{B}^{\mathrm{op}}, \mathcal{E}]$, and hence also in $\mathbf{Sh}_{\mathcal{E}}(\mathbb{B}, P)$.

Now let $f: \mathcal{F} \to \mathcal{E}$ be an arbitrary tidy morphism, and let

$$\mathcal{F} \xrightarrow{\ h\ } \mathcal{G} \xrightarrow{\ g\ } \mathcal{E}$$

be its pure–entire factorization. Suppose given a filtered internal category $\mathbb{C}$ in $\mathcal{G}$; then by 3.4.17 the category $g_*\mathbb{C}$ is filtered in $\mathcal{E}$, and since filtered categories are necessarily well-supported the counit map $g^* g_* \mathbb{C} \to \mathbb{C}$ is epic, and hence final in the sense of B2.5.12. Hence the diagram

$$\begin{tikzcd}
{[\mathbb{C}, \mathcal{G}]} \arrow[r, "\varinjlim_{\mathbb{C}}"] \arrow[d] & \mathcal{G} \\
{[g^* g_* \mathbb{C}, \mathcal{G}]} \arrow[ur, "\varinjlim_{g^* g_* \mathbb{C}}"'] &
\end{tikzcd}$$

commutes, where the vertical arrow is induced by pullback along the counit. Now consider the diagram

$$\begin{tikzcd}
{[h^*\mathbb{C}, \mathcal{F}]} \arrow[r] \arrow[d] & {[\mathbb{C}, \mathcal{G}]} \arrow[d] & \\
{[h^* g^* g_* \mathbb{C}, \mathcal{F}]} \arrow[r] \arrow[d, "\varinjlim"] & {[g^* g_* \mathbb{C}, \mathcal{G}]} \arrow[r] \arrow[d, "\varinjlim"] & {[g_* \mathbb{C}, \mathcal{E}]} \arrow[d, "\varinjlim"] \\
\mathcal{F} \arrow[r, "h_*"] & \mathcal{G} \arrow[r, "g_*"] & \mathcal{E}
\end{tikzcd}$$

where the horizontal arrows are all direct image functors. If we chase a discrete opfibration $\mathbb{F} \to h^*\mathbb{C}$ around the upper square, we obtain the pullbacks of $h_*\mathbb{F}$ either way around the square

$$\begin{tikzcd}
g^*g_*\mathbb{C} \arrow[r] \arrow[d] & \mathbb{C} \arrow[d] \\
h_*h^*g^*g_*\mathbb{C} \arrow[r] & h_*h^*\mathbb{C}
\end{tikzcd}$$

formed by the unit and counit maps; but this square commutes by naturality of the unit, and so the top square in the large diagram commutes up to isomorphism. Also, the right lower square and the outside of the lower rectangle commute up to isomorphism, since gh and g are both tidy. So we conclude that, for any $\mathbb{F} \to h^*\mathbb{C}$, the Beck–Chevalley morphism $\varinjlim_{\mathbb{C}}[\mathbb{C}, h]_*(\mathbb{F}) \to h_* \varinjlim_{h^*\mathbb{C}}(\mathbb{F})$ is mapped by g_* to an isomorphism; i.e., considering it as a map between sheaves on a Boolean algebra $\mathbb{B}$ in $\mathcal{E}$, it is 'bijective on global sections'. Now each element b of B corresponds (via the principal ideal $\downarrow(b)$) to a clopen subtopos of $\mathcal{G}$, and hence (via pullback) to a clopen subtopos $\mathcal{F}'$ of $\mathcal{F}$; the composite $\mathcal{F}' \to \mathcal{F} \to \mathcal{E}$ is still proper, so on restricting to it we may similarly deduce that the Beck–Chevalley morphism is bijective on sections over b, for any b; i.e. it is an isomorphism in $\mathbf{Sh}(\mathbb{B}, P) = \mathcal{G}$.

We are still not finished; for we have to consider filtered categories in arbitrary slices of $\mathcal{G}$, not just in $\mathcal{G}$ itself. However, these can be dealt with in a similar manner, using the fact that every object of $\mathcal{G}$ is an $\mathcal{E}$-indexed colimit of complemented subterminal objects. We shall not go into the details here.

For geometric morphisms over a base topos $\mathcal{S}$, local connectedness may be viewed as a strengthening of the property of being $\mathcal{S}$-essential. In the same way, tidiness of a morphism $f: \mathcal{F} \to \mathcal{E}$ may be seen as a strengthening of the 'relative' condition that the direct image functor f_* should preserve filtered $\mathcal{S}$-indexed colimits (that is, the special case of Definition 3.4.2 where the object A and the filtered category $\mathbb{C}$ are assumed to come from $\mathcal{S}$ and $\mathcal{S}/A$ respectively). This relative notion of tidiness has also been studied by Moerdijk and Vermeulen [858]: in particular, they proved the dual of Pitts's Theorem 2.3.17, which we state here (although we refer to [858] for the proof).

Theorem 3.4.18 *Suppose given a comma square*

$$\begin{tikzcd}
(f \downarrow g) \arrow[r, "v"] \arrow[d, "u"'] \arrow[dr, phantom, "\to"] & \mathcal{G} \arrow[d, "g"] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$, where $\mathcal{S}$ is a topos with a natural number object. If g is tidy relative to $\mathcal{S}$, then u is tidy, and the Beck–Chevalley natural transformation $f^*g_* \to u_*v^*$ is an isomorphism. □

Suggestions for further reading: Johnstone [519], Moerdijk & Vermeulen [857, 858].

C3.5 Atomic morphisms

In this section we study those geometric morphisms whose inverse image functors are logical. Of course, we already know quite a bit about such morphisms: we saw in A2.3.2 that every local homeomorphism has logical inverse image, in A2.2.10(ii) that any such morphism must be essential, and in A2.4.8 that the left adjoint of its inverse image functor must preserve monomorphisms. Also, such a morphism must be locally connected, since its inverse image commutes with the functors Π_f.

Traditionally, a geometric morphism with logical inverse image is said to be *atomic*. This is perhaps not a particularly good name, since it fails to suggest that the condition is a strengthening of 'open' and of 'locally connected'; in fact it derives from a characteristic property of atomic **Set**-toposes (that is, toposes which possess an atomic geometric morphism to **Set**), as we shall see in 3.5.9 below, and is not really appropriate for morphisms between non-Boolean toposes. Nevertheless, it is by now a well-established term, and we have not attempted to depart from it.

Lemma 3.5.1

(i) *A composite of atomic morphisms is atomic.*

(ii) *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

be a commutative diagram of geometric morphisms, where h is atomic and g is an open surjection. Then f is atomic.

Proof

(i) is trivial.

(ii) is proved in exactly the same way as 3.3.2(iv), but using the power-object comparison map $\phi_A\colon f^*(PA) \to P(f^*A)$ instead of the exponential comparison map. □

Corollary 3.5.2 *Let $\mathcal{E}$ be a Grothendieck topos with enough points. Then $\mathcal{E}$ is Boolean iff the unique geometric morphism $\gamma\colon \mathcal{E} \to \mathbf{Set}$ is atomic.*

Proof Clearly, if γ is atomic then $\mathcal{E}$ must be Boolean, since γ^* preserves the isomorphism $(1 \amalg 1) \cong \Omega$. Conversely, if $\mathcal{E}$ is Boolean, we apply 3.5.1(ii) to the diagram

$$\begin{tikzcd} \mathbf{Set}/K \arrow[r, "g"] \arrow[dr] & \mathcal{E} \arrow[d, "\gamma"] \\ & \mathbf{Set} \end{tikzcd}$$

where K is a sufficient set of points of $\mathcal{E}$ (cf. 2.2.12); note that g is open by the remark before 3.1.9. $\square$

For Grothendieck toposes without enough points, atomicity over **Set** is a stronger condition than Booleanness. Indeed, for localic toposes we have:

Lemma 3.5.3 *Let X be a locale. Then the following are equivalent:*

(i) *X is discrete (that is, $\mathcal{O}(X)$ is a full power-set).*

(ii) *$\mathbf{Sh}(X)$ is atomic over* **Set**.

(iii) *$\mathbf{Sh}(X)$ is Boolean and has enough points.*

Proof (i) $\Rightarrow$ (iii) follows from the fact that if X is discrete we have $\mathbf{Sh}(X) \simeq \mathbf{LH}/X \cong \mathbf{Set}/X$.

(iii) $\Rightarrow$ (ii) follows from 3.5.2.

(ii) $\Rightarrow$ (i): Let $\gamma_!\colon \mathbf{Sh}(X) \to \mathbf{Set}$ be the left adjoint of the inverse image functor. Then by A2.4.8 we have $\mathcal{O}(X) \cong \mathrm{Sub}_{\mathbf{Sh}(X)}(1) \cong \mathrm{Sub}_{\mathbf{Set}}(\gamma_!(1))$. $\square$

An atomic **Set**-topos which is not localic need not have enough points: we give a counterexample in D3.4.14. However, one part of the proof of 3.5.3 survives in general: if $\gamma\colon \mathcal{E} \to \mathbf{Set}$ is atomic, then for every object A of $\mathcal{E}$ we have $\mathrm{Sub}_{\mathcal{E}}(A) \cong \mathrm{Sub}_{\mathbf{Set}}(\gamma_!(A))$, so subobject lattices in $\mathcal{E}$ are full power-sets (that is, complete atomic Boolean algebras). This is the origin of the name 'atomic topos', as hinted earlier.

The equivalence of (i) and (ii) in 3.5.3 remains valid over an arbitrary base:

Lemma 3.5.4

(i) *A connected atomic morphism is hyperconnected.*

(ii) *If $f\colon \mathcal{F} \to \mathcal{E}$ is atomic, then the factorization $\mathcal{F} \to \mathcal{E}/f_!1 \to \mathcal{E}$ of* 3.3.5(i) *coincides with the hyperconnected–localic factorization of f, and both halves of it are atomic.*

(iii) *A localic morphism is atomic iff it is a local homeomorphism.*

Proof (i) If $f\colon \mathcal{F} \to \mathcal{E}$ is atomic and connected, then for every object A of $\mathcal{E}$ we have a bijection $\mathrm{Sub}(f^*A) \to \mathrm{Sub}(f_!f^*A) \cong \mathrm{Sub}(A)$ induced by $f_!$, whose

inverse must be induced by f^*. In particular, every subobject of f^*A is in the image of f^*; i.e. f satisfies condition (ii) of A4.6.6.

(ii) As in the proof of 3.3.5(i), we may factor the inverse image functor $\mathcal{E}/f_!1 \to \mathcal{F}$ as

$$\mathcal{E}/f_!1 \xrightarrow{(f/f_!1)^*} \mathcal{F}/f^*f_!1 \xrightarrow{\eta_1^*} \mathcal{F}$$

where η is the unit of $(f_! \dashv f^*)$; hence it is logical. So by (i) the morphism $\mathcal{F} \to \mathcal{E}/f_!1$ is hyperconnected; and since the other factor is localic by A4.6.2(a), the factorization is the hyperconnected–localic one.

(iii) is immediate from (ii), since if f is localic then the first half of the factorization in (ii) is an equivalence. $\square$

For dealing with atomic morphisms over an arbitrary base, the following characterization is often useful. For the moment, let us write $D\colon \mathcal{E} \to \mathbf{Loc}(\mathcal{E})$ for the functor which sends an object A to the discrete locale on A (defined by $\mathcal{O}(DA) = PA$), and let $X \mapsto X_p$ denote its right adjoint which sends a locale X to its object of points (i.e. the hom-object $\mathbf{Frm}(\mathcal{E})\,(\mathcal{O}(X), \Omega)$ of the locally internal category $\mathbb{F}\mathsf{rm}(\mathcal{E})$; cf. 1.2.2).

Lemma 3.5.5 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is atomic.*

(ii) *For any internal locale X in $\mathcal{E}$, we have $f^*(X_p) \cong (f^*X)_p$.*

(iii) *The functor $f_!\colon \mathbf{Loc}(\mathcal{F}) \to \mathbf{Loc}(\mathcal{E})$ preserves discrete locales.*

(iv) *For every object B of $\mathcal{F}$, the composite $\mathcal{F}/B \to \mathcal{F} \to \mathcal{E}$ can be factored as a hyperconnected morphism followed by a local homeomorphism.*

Proof (i) $\Rightarrow$ (iv) follows immediately from 3.5.4(ii).

(iv) $\Rightarrow$ (iii): By the definition of $f_!$ on locales in Section C2.4, $f_!(DB)$ is the internal locale in $\mathcal{E}$ corresponding to the localic part of the composite $\mathbf{Sh}_{\mathcal{F}}(DB) \simeq \mathcal{F}/B \to \mathcal{F} \to \mathcal{E}$. Condition (iv) says that this locale is discrete.

(iii) $\Rightarrow$ (ii): Let X be an internal locale in $\mathcal{E}$. Then, for any object B of $\mathcal{F}$, we have bijections

$$\begin{array}{rcll}
B & \longrightarrow & (f^*X)_p & \\ \hline
DB & \longrightarrow & f^*X & \\ \hline
f_!(DB) & \longrightarrow & X & \\ \hline
f_!(DB) & \longrightarrow & D(X_p) & \text{since } f_!(DB) \text{ is discrete} \\ \hline
DB & \longrightarrow & f^*(D(X_p)) & \\ \hline
DB & \longrightarrow & D(f^*(X_p)) & \text{since } f^* \text{ preserves discrete locales} \\ \hline
B & \longrightarrow & f^*(X_p) &
\end{array}$$

so $(f^*X)_p \cong f^*(X_p)$.

(ii)$\Rightarrow$(i): Given an object A of $\mathcal{E}$, let S^A be the A-fold power of the Sierpiński locale, i.e. the locale such that $\mathcal{O}(S^A)$ is the free frame FA generated by A (cf. 1.1.4). From the adjunction $(f^\# \dashv f_*)$, it is clear that $f^\#(FA) \cong F(f^*A)$, so on taking objects of points we obtain

$$f^*(PA) \cong f^*((S^A)_p) \cong (f^*(S^A))_p \cong (S^{f^*A})_p \cong P(f^*A) \ ,$$

i.e. f^* is logical. $\square$

Condition (iv) of 3.5.5 should be compared with 3.3.5(ii). In connection with 3.5.5(ii), we remark that if f is atomic then the functor $f^\#: \mathbf{Frm}(\mathcal{E}) \to \mathbf{Frm}(\mathcal{F})$ *is* simply the functor f^* applied to internal frames in $\mathcal{E}$; the fact that logical functors preserve completeness of internal posets was proved in B2.3.10(i). Using this, it is not hard to give a direct proof of the implication (i)$\Rightarrow$(ii) in 3.5.5 (as is done, for example, in [560]).

An interesting feature of atomic morphisms is that their one-sided inverses, if any, are necessarily open.

Proposition 3.5.6 *Let $f: \mathcal{F} \to \mathcal{E}$ be an atomic morphism, let A be an object of $\mathcal{E}$ and let $g: \mathcal{E}/A \to \mathcal{F}$ be a geometric morphism over $\mathcal{E}$. Then*

(i) *g is open;*

(ii) *if f is connected and $A \to 1$ is epic, g is surjective.*

Proof It is convenient to prove the second assertion first.

(ii) Since both f_* and f^* preserve Ω, we know that their left adjoints f^* and $f_!$ both induce bijections on subobject lattices, by A4.6.6(vi) and A2.4.8. Moreover, an easy diagram-chase shows that, for any $B \in \text{ob}\, \mathcal{F}$, the composite bijection $\text{Sub}(B) \to \text{Sub}(f^*f_!B)$ is inverse to the operation of pulling back along the unit map $B \to f^*f_!B$; so, by considering the image of the latter, we see that it must be an epimorphism. Now, for any $B' \rightarrowtail B$, we have a pullback square

$$\begin{tikzcd} B' \arrow[r, two heads] \arrow[d, tail] & f^*f_!B' \arrow[d, tail] \\ B \arrow[r, two heads] & f^*f_!B \end{tikzcd} \quad ;$$

applying g^* to this, we obtain a pullback square

$$\begin{tikzcd} g^*B' \arrow[r, two heads] \arrow[d, tail] & A^*f_!B' \arrow[d, tail] \\ g^*B \arrow[r, two heads] & A^*f_!B \end{tikzcd} \quad ,$$

since $g^*f^* \cong A^* : \mathcal{E} \to \mathcal{E}/A$. Hence g^*, like $f_!$, preserves properness of subobjects; but since it also (unlike $f_!$ – cf. A2.3.8) preserves equalizers, it is conservative by A1.2.4.

(i) First we observe that we may reduce to the case $A = 1$: for we have a pullback

$$\begin{tikzcd} \mathcal{F}/f^*A \arrow[r] \arrow[d, "f/A"] & \mathcal{F} \arrow[d, "f"] \\ \mathcal{E}/A \arrow[r] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{Top}$, where f/A is atomic by A2.3.7, and thus g induces a morphism $\mathcal{E}/A \to \mathcal{F}/f^*A$ over $\mathcal{E}/A$, whose openness implies that of g itself by 3.2.23.

Let B be an object of $\mathcal{F}$, and suppose we are given a subobject $C \rightarrowtail g^*B$. As in the proof of (ii), it is easy to prove that for any $B' \rightarrowtail B$ the naturality square

$$\begin{tikzcd} B' \arrow[r] \arrow[d, tail] & f^*f_!B' \arrow[d, tail] \\ B \arrow[r] & f^*f_!B \end{tikzcd}$$

is a pullback (by considering the classifying maps of the two vertical morphisms), although its horizontal edges are no longer necessarily epic. Applying g^* to this square, we obtain a pullback

$$\begin{tikzcd} g^*B' \arrow[r] \arrow[d, tail] & f_!B' \arrow[d, tail] \\ g^*B \arrow[r] & f_!B \end{tikzcd} \quad ;$$

so if we take $B' \rightarrowtail B$ to be the unique subobject such that $f_!B'$ is the image of the composite $C \rightarrowtail g^*B \to f_!B$, we obtain the unique smallest subobject of B such that $C' \leq g^*B'$.

Now by 3.5.5(iv) we have a connected atomic morphism $\mathcal{F}/B' \to \mathcal{E}/f_!B'$. Also, the composite $\mathcal{E}/C \to \mathcal{E}/g^*B' \to \mathcal{F}/B'$ is a morphism over $\mathcal{E}/f_!B'$; and since $C \to f_!B'$ is epic by the definition of B', it follows from (ii) that this latter composite is a surjection. So the image of $\mathcal{E}/C \to \mathcal{E}/g^*B \to \mathcal{F}/B$ (in the sense of the surjection–inclusion factorization) is $\mathcal{F}/B'$; thus we have verified condition (vi) of 3.1.7 for g. □

Remark 3.5.7 Using the descent theorem for open surjections, we shall see in 5.2.12 below that if $f: \mathcal{E} \to \mathcal{S}$ is connected atomic and has a section, then $\mathcal{E}$ may be identified with the topos of continuous G-sets for a suitable localic group G in $\mathcal{S}$. In particular, it is bounded over $\mathcal{S}$. That is, *a connected atomic morphism with a section is necessarily bounded.* The assumption of having a section is essential here; recall that two of our three prime examples of unbounded toposes over **Set** (cf. B3.1.14) are the topos $\mathbf{Unif}(G)$ where G is a topological group having no smallest open subgroup as in A2.1.7, and the topos $[\mathcal{G}, \mathbf{Set}]$ where $\mathcal{G}$ is a 'large' group as in A2.1.4. Both of these are easily seen to be connected and atomic, but neither has a point. (For both toposes, the forgetful functor to **Set** preserves finite limits and all colimits which exist, but fails to have a right adjoint – in one case because the topos fails to be cocomplete, and in the other because the solution-set condition fails.)

Our next task is, as usual, to give a site characterization of bounded atomic morphisms. Since atomic morphisms are locally connected, we know that for any bounded atomic morphism $p: \mathcal{E} \to \mathcal{S}$ we can choose a site of definition for $\mathcal{E}$ consisting of objects A which are connected, i.e. satisfy $p_!(A) \cong 1$. In the case when p is atomic, connected objects of $\mathcal{E}$ are generally called *atoms*; by 3.5.5(iii), an atom A has the property that $p_*(\Omega_{\mathcal{E}}^A) \cong \Omega_{\mathcal{S}}$, and in particular that the map $\mathrm{Sub}_{\mathcal{S}}(1) \to \mathrm{Sub}_{\mathcal{E}}(A)$ sending U to $A \times p^*U$ is bijective. (The name 'atom' is really appropriate only in the classical case $\mathcal{S} = \mathbf{Set}$, when this condition says that A has no subobjects other than itself and 0 (and these two are distinct); but, as we remarked earlier, it seems hard to avoid using it in the more general context.)

We note that any morphism between atoms is an epimorphism in $\mathcal{E}$; for if $f: A \to B$ had image $B \times p^*U$, then we should have $A \cong A \times p^*U$. In particular, therefore, if $\mathbb{C}$ is any internal full subcategory of $\mathbb{E}$ (over $\mathcal{S}$) on a generating set of atoms, each morphism of $\mathbb{C}$ generates a covering sieve in the induced coverage on $\mathbb{C}$. We note also that $\mathbb{C}$ satisfies the right Ore condition (so that, by A2.1.11(h), these sieves do indeed form a coverage): if $f: A \to C$ and $g: B \to C$ are two morphisms of $\mathbb{C}$ with common codomain, then the pullback $A \times_C B$ in $\mathcal{E}$ is nonzero, and so admits a morphism from some object of $\mathbb{C}$.

In 2.1.12(c) we considered two further conditions on a site $(\mathcal{C}, T)$ as above, designed to ensure that T coincides with the canonical coverage on $\mathcal{C}$: namely, that every morphism of $\mathcal{C}$ is an effective epimorphism (i.e., if $g: A \to C$ coequalizes every pair of morphisms coequalized by $f: A \to B$, then g factors uniquely through f) and that $\mathcal{C}$ does not have a strict initial object. The first of these is true in general (since, for any$f: A \to B$, the family of morphisms from objects of $\mathbb{C}$ to the kernel-pair $A \times_B A$ is epimorphic); but the second is special to the case $\mathcal{S} = \mathbf{Set}$ (and even in that case it fails when the only atom is the terminal object 1, i.e. when $\mathcal{E} = \mathbf{Set}$). Also, the first condition is not stable under change of base; so we do not include it in our definition of an 'atomic site', even though it was part of Barr's original definition [81].

One further feature of the case $\mathcal{S} = \mathbf{Set}$ deserves mention. In a bounded atomic **Set**-topos there are, up to isomorphism, only a set of atoms, since if $\{G_i \mid i \in I\}$ is any generating set then each atom must be a surjective image of a single G_i. Conversely, any epimorphic image of an atom is an atom; so, by taking an arbitrary generating subcategory of atoms and closing it under quotients in $\mathcal{E}$, we may assume that our site contains (up to isomorphism) all the atoms of $\mathcal{E}$. However, this property is not needed for our characterization.

Theorem 3.5.8 *For a bounded geometric morphism $p: \mathcal{E} \to \mathcal{S}$, the following are equivalent:*

(i) *p is atomic.*

(ii) *There exists an internal site $(\mathbb{C}, T)$ for $\mathcal{E}$ over $\mathcal{S}$ such that $\mathbb{C}$ satisfies the right Ore condition, and T consists either of all inhabited sieves on objects of $\mathbb{C}$ or of all sieves generated by single morphisms of $\mathbb{C}$.*

(iii) *There exists a site $(\mathbb{C}, T)$ as in* (ii), *but with the additional property that each morphism of $\mathbb{C}$ is an effective epimorphism.*

Proof (i) $\Rightarrow$ (iii) is proved in the discussion above, and (iii) $\Rightarrow$ (ii) is trivial.

For (ii) $\Rightarrow$ (i), we note first that a site as in (ii) is locally connected, as we saw in 3.3.11(b); so we already know that p^* is a cartesian closed functor, and we need only show that it preserves Ω. For this, the argument is similar to (but easier than) that of 3.1.19(i): recall from the proof of 3.3.10 that p^* simply sends $\Omega_{\mathcal{S}}$ to the constant diagram on $\mathbb{C}^{\mathrm{op}}$ with value $\Omega_{\mathcal{S}}$, so we have only to verify that, for each object U of $\mathbb{C}$, the object $\mathrm{Sv}_T(U)$ of T-closed sieves on U, as described in the proof of 3.1.19(i), is isomorphic to Ω. But this is easy, since a T-closed sieve R is entirely determined by the truth-value of the assertion 'R is inhabited': since every inhabited sieve is T-covering (or at least contains a T-covering sieve), to the extent that R is inhabited it must contain all morphisms with codomain U. (Also, because all T-covering sieves are inhabited, the sieves

$$\{a: C_1 \mid d_0(a) = U \wedge p\}$$

(where $p: \Omega_{\mathcal{S}}$) are all T-closed.) □

We note that 3.5.8 has a straightforward strengthening, similar to those for our site characterizations of openness and local connectedness: p is connected and atomic iff the site in (ii) or (iii) may be chosen to have a terminal object.

Before going on to prove the pullback-stability and inverse limit theorems, we pause to discuss some important examples of atomic sites in the case $\mathcal{S} = \mathbf{Set}$.

Examples 3.5.9 (a) Any group G (more generally, any small groupoid) is an atomic site when equipped with the coverage in which only the maximal sieves cover (as we noted in A2.1.11(i), these are the only inhabited sieves in this case). So $[G, \mathbf{Set}]$ is an atomic topos (as we already knew from A2.1.4). The atoms of $[G, \mathbf{Set}]$ are the transitive G-sets (equivalently, the sets $G : H$ of left cosets of

arbitrary subgroups H of G); so we may obtain an alternative site of definition for $[G, \mathbf{Set}]$ by taking the full subcategory of all such G-sets, with the coverage in which every morphism generates a covering sieve. (However, we note that in this case the coverage is rigid in the sense of 2.2.18, since there is a largest atom, namely G itself; so we could alternatively use the Comparison Lemma 2.2.3 to derive the equivalence of the topos of sheaves on this site with $[G, \mathbf{Set}]$.)

(*b*) For a topological group G, the atoms of $\mathbf{Cont}(G)$ are again the transitive continuous G-sets, i.e. the G-sets $G : H$ where H is an open subgroup. So we may take these as the objects of our atomic site; equivalently, as we observed in 2.2.4(*c*), we may take the objects of the site to be the open subgroups themselves, with morphisms $H \to K$ labelled by those cosets gK of K for which $H \subseteq gKg^{-1}$, and the composite of $gK\colon H \to K$ and $hL\colon K \to L$ defined to be ghL.

(*c*) As we remarked in 2.1.12(*c*), the category $\mathbf{Set}_{fe}$ of finite sets and surjections, and the dual of the category $\mathbf{Set}_{fm}$ of finite sets and injections, are both atomic sites when equipped with the coverage consisting of all sieves generated by single morphisms. (In fact they are both special cases of the previous example, for suitable topological groups G; we prove this explicitly for $\mathbf{Set}_{fm}^{\mathrm{op}}$ in D3.4.10, and for $\mathbf{Set}_{fe}$ in D3.4.12. We recall also that in A2.1.11(*h*) we showed that the sheaves for the atomic coverage on $\mathbf{Set}_{fm}^{\mathrm{op}}$ are exactly the pullback-preserving functors; so we can now conclude that *the category of pullback-preserving functors* $\mathbf{Set}_{fm} \to \mathbf{Set}$ *is an atomic topos.*

(*d*) A similar argument shows that if $\mathbf{Ord}_{fm}$ denotes the category of finite totally ordered sets and order-preserving injections, then the pullback-preserving functors $\mathbf{Ord}_{fm} \to \mathbf{Set}$ are exactly the sheaves for the atomic coverage on $\mathbf{Ord}_{fm}^{\mathrm{op}}$, so they form an atomic topos. Once again, we shall meet this topos again in Section D3.4, where we shall see that it too can be identified with a topos of continuous G-sets. For more information about this topos, see [529].

From 3.5.8, we may obtain a characterization of connected atomic morphisms in terms of fibrations of sites. In this context, we shall say that a functor $F\colon \mathcal{C} \to \mathcal{D}$ between the underlying categories of two sites *creates covers* if, given any sieve R on an object U of $\mathcal{C}$, R is covering iff the sieve generated by $\{Fa \mid a \in R\}$ covers FU in $\mathcal{D}$; note that this implies that F both preserves and reflects covers as defined in Section C2.3, but it is stronger than either of these conditions.

Proposition 3.5.10 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be a geometric morphism between bounded $\mathcal{S}$-toposes. Then f is connected and atomic iff it can be defined by a fibration of sites $(P, T)\colon (\mathbb{D}, K) \to (\mathbb{C}, J)$ in $\mathcal{S}$ such that P creates covers.*

Proof First suppose f is connected and atomic. Choose any site $(\mathbb{C}, J)$ for $\mathcal{E}$ over $\mathcal{S}$; let $(\mathbb{B}, H)$ be an atomic site (with a terminal object) for $\mathcal{F}$ over $\mathcal{E}$, and externalize it as in 2.5.7 to obtain a continuous fibration. The assertion that the H-covering sieves are exactly the inhabited ones means that we can simplify the description of the coverage $K = J \rtimes H$ which we obtain on $\mathbb{D} = \mathbb{C} \rtimes \mathbb{B}$: a

sieve S on (U, V) is covering iff the set of all $a: U' \to U$ for which there exists some b with $(a, b): (U', V') \to (U, V) \in S$ is J-covering. But this is exactly the statement that P creates covers.

Conversely, suppose P creates covers. We claim first that (P, T) satisfies the condition of 3.3.12; in fact, if we have a K-covering sieve S on V and a commutative square

$$\begin{tikzcd} U \arrow[r, "a_2"] \arrow[d, "a_1"] & PV_2 \arrow[d, "Pb_2"] \\ PV_1 \arrow[r, "Pb_1"] & PV \end{tikzcd}$$

with $b_1, b_2 \in S$, then U can be J-covered by morphisms $c: PV_3 \to U$ for which the composites $a_i c$ $(i = 1, 2)$ are equal to Pd_i for some morphisms d_i making the square

$$\begin{tikzcd} V_3 \arrow[r, "d_2"] \arrow[d, "d_1"] & V_2 \arrow[d, "b_2"] \\ V_1 \arrow[r, "b_1"] & V \end{tikzcd}$$

commute in $\mathbb{D}$. To prove this, choose prone liftings of a_1, a_2 and the diagonal of the first square above, to obtain a diagram

$$\begin{tikzcd} \widetilde{U_1} \arrow[r, "e_1"] \arrow[d] & \widetilde{U} \arrow[d] & \widetilde{U_2} \arrow[l, "e_2"'] \arrow[d] \\ V_1 \arrow[r, "b_1"] & V & V_2 \arrow[l, "b_2"'] \end{tikzcd}$$

where e_1 and e_2 are vertical (that is, their images under P are the identity on U). Since P creates covers, each of e_1 and e_2 generates a K-covering sieve; hence, by the remark after 2.1.8, the intersection of these two sieves (i.e. the sieve of all morphisms which factor through both e_1 and e_2) is also K-covering. The image of this sieve under P is the cover of U which we seek.

We therefore know that f is connected and locally connected, so it remains to show that f^* preserves Ω. But we may do this by direct calculation, as in the proof of 3.1.21: we have to show that the mapping $\phi_V : \Omega_{\mathcal{E}}(PV) \to \Omega_{\mathcal{F}}(V)$, which sends a J-closed sieve R on PV to the K-closed sieve $\{b: V' \to V \mid Pb \in R\}$, is bijective. But it is easy to see that a K-closed sieve S is generated by the prone morphisms which it contains, since (as we already observed) every vertical

morphism in $\mathbb{D}$ must generate a K-covering sieve, and hence we may uniquely recover R from S, as the sieve of all morphisms $U' \to PV$ whose prone liftings lie in S. □

Corollary 3.5.11 *Let*

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

be an inverse sequence in $\mathfrak{BTop}/\mathcal{S}$*, where each* f_i *is connected and atomic. Then the legs* $f_\infty^n : \mathcal{E}_\infty \to \mathcal{E}_n$ *of the limit cone in* $\mathfrak{BTop}/\mathcal{S}$ *are connected and atomic.*

Proof As usual, we represent the f_n by an inverse sequence of fibrations of sites satisfying the condition of 3.5.10. We must show that the projections $P_\infty^n : \mathbb{C}_\infty \to \mathbb{C}_n$ (in the notation of 2.5.11) create covers; but we already know that they preserve them, by 3.1.22. Suppose given a sieve S on an object $(U_m \mid m \geq 0)$ of $\mathbb{C}_\infty$, whose image under P_∞^n generates a J_n-covering sieve. Then, for any $n' \geq n$, its image under $P_\infty^{n'}$ must generate a $J_{n'}$-covering sieve, since $P_{n'}^n$ creates covers; and we can choose n' so large that $U_{j+1} \cong T_{j+1}U_j$ for all $j \geq n'$, i.e. so that $(U_m \mid m \geq 0) \cong T_\infty^{n'}(U_{n'})$. But now $T_\infty^{n'}$ preserves covers; so on applying it to our covering sieve on $U_{n'}$ we get a covering sieve generated by the $P_\infty^{n'}$-prone parts of all the morphisms in S. Hence, to prove that S is covering, it suffices by the local character axiom (L) to verify that the sieve generated by a single $P_\infty^{n'}$-vertical morphism is covering. But this follows by a similar argument: the morphism in question is in the image of $T_\infty^{n''}$ for some $n'' \geq n'$, and the morphism in $\mathbb{C}_{n''}$ which maps to it is $P_{n'}^{n''}$-vertical and hence generates a covering sieve there. □

We are now ready to prove the pullback-stability theorem for atomic morphisms. As usual, it has two forms, depending on which side of the pullback square is assumed to be bounded.

Theorem 3.5.12 *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$*, where* f *is atomic and either* f *or* g *is bounded. Then* k *is atomic.*

Proof (i) In the case when f is bounded, we may represent $\mathcal{F}$ as $\mathbf{Sh}_{\mathcal{E}}(\mathbb{C}, T)$, where $(\mathbb{C}, T)$ is a site as in 3.5.8(ii), and then $\mathcal{H} \simeq \mathbf{Sh}_{\mathcal{G}}(g^*\mathbb{C}, g^\# T)$. Now $g^*\mathbb{C}$ clearly inherits the right Ore condition from $\mathbb{C}$, since this may be expressed by a regular sequent (it is part of the assertion that $\mathbb{C}^{\mathrm{op}}$ is weakly filtered,

cf. B2.6.2). The condition that T consists of all inhabited sieves is not in general preserved under the passage to $g^{\#}T$, since the comparison map $g^*(P^+C_1) \to P^+(g^*C_1)$ need not be epic; but the condition that T consists precisely of the sieves generated by single morphisms of $\mathbb{C}$ clearly is preserved. So the result is immediate from 3.5.8.

(ii) In the case when g is bounded, we as usual consider the two sub-cases when g has the form $[\mathbb{C}, \mathcal{E}] \to \mathcal{E}$ and when it is an inclusion. For the first of these, we recall from the proof of 3.1.24 that it is possible to give an explicit construction in terms of $\mathcal{E}$ of the power object of an internal diagram F on $\mathbb{C}$; this construction uses only structure preserved by f^*, and so it is itself preserved by $[\mathbb{C}, f]^*$, i.e. the latter is logical. For the case when g is the inclusion of a subtopos $\mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$, we note that by 3.1.25 we have a pullback

$$\begin{tikzcd}
f^*(\Omega_j) \arrow[r] \arrow[d, tail] & \Omega_{j'} \arrow[d, tail] \\
f^*(\Omega) \arrow[r, "\phi_1"] & \Omega
\end{tikzcd}$$

where j' is the pullback local operator on $\mathcal{F}$; but ϕ_1 is an isomorphism in this case, and so the top edge of the pullback square is also an isomorphism, i.e. k^* (which is the restriction of f^* to sheaves, by the second proof of 3.3.15) preserves Ω. But it is also a cartesian closed functor by 3.3.15(ii), and hence logical. □

Remark 3.5.13 An alternative proof of the inclusion case of 3.5.12(ii) may be given using the criterion of 3.5.5(iv). We know that the pullback k is locally connected and hence essential, by 3.3.15; now, given any object B of $\mathcal{H}$, we set $A = h_*B$ and consider the cube

$$\begin{tikzcd}
\mathcal{H}/B \arrow[rr] \arrow[dr] \arrow[dd] & & \mathcal{G}/k_!B \arrow[dr] \arrow[dd] & \\
& \mathcal{H} \arrow[rr, "k"] \arrow[dd, "h"] & & \mathcal{G} \arrow[dd, "g"] \\
\mathcal{F}/A \arrow[rr] \arrow[dr] & & \mathcal{E}/f_!A \arrow[dr] & \\
& \mathcal{F} \arrow[rr, "f"] & & \mathcal{E}
\end{tikzcd} .$$

Here the left vertical face is a pullback since h is an inclusion and so $h^*A \cong B$, and the right vertical face is a pullback since the Beck–Chevalley condition for locally connected morphisms yields $g^*f_!A \cong k_!h^*A \cong k_!B$. Since the front face is a pullback by assumption, the back face is too; hence from the fact that $\mathcal{F}/A \to \mathcal{E}/f_!A$ is hyperconnected we may deduce that $\mathcal{H}/B \to \mathcal{G}/k_!B$ is hyperconnected. So k satisfies the condition of 3.5.5(iv).

We may now deduce another characterization of (bounded) atomic morphisms, paralleling the characterization of discrete locales which we gave in 3.1.15.

Corollary 3.5.14 *A bounded geometric morphism $f: \mathcal{F} \to \mathcal{E}$ is atomic iff both f and the diagonal morphism $\mathcal{F} \to \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ are open.*

Proof First suppose f is atomic. Then f is certainly open, and the projections $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \rightrightarrows \mathcal{F}$ are both atomic by 3.5.12. But the diagonal is a section of either of these, so it is open by 3.5.6.

For the converse, we note first that the property of having open diagonal is stable under composition of geometric morphisms, and is inherited by both halves of the hyperconnected–localic factorization, by arguments exactly like those which we gave in the 'proper' case in 3.2.25. Hence in particular, for any object B of $\mathcal{F}$, the localic part of the composite $\mathcal{F}/B \to \mathcal{F} \to \mathcal{E}$ must be open with open diagonal; so by 3.1.15 it corresponds to a discrete internal locale in $\mathcal{E}$. So we have verified the condition of 3.5.5(iii), and f is atomic. □

Suggestions for further reading: Barr & Diaconescu [81], Johnstone [541], Joyal & Tierney [560].

C3.6 Local maps

In Section C3.3, we studied those geometric morphisms $f: \mathcal{F} \to \mathcal{E}$ for which f^* has an $\mathcal{E}$-indexed left adjoint; in the present section our concern is with those f for which f_* has an $\mathcal{E}$-indexed right adjoint. At this point, the 'duality' between the odd- and even-numbered sections of this chapter finally breaks down: by the characterization of bounded atomic morphisms in 3.5.14, their 'proper' dual is the class of proper separated maps, which neither contains nor is contained in the class of local maps which we study in this section. In fact, as we shall see, the 'open' dual of the notion of local map is a different strengthening of the notion of locally connected morphism, which we shall characterize in 3.6.16 below under the name 'totally connected morphism'.

Of course, f_* has a right adjoint iff it is the inverse image of a geometric morphism $c: \mathcal{E} \to \mathcal{F}$, iff f has a *left* adjoint c in the 2-category $\mathfrak{Top}$. But this condition alone will not ensure that the adjunction $(c^* \dashv c_*)$ is indexed over $\mathcal{E}$ (though it will, of course, be indexed over $\mathcal{F}$); so our first concern is to determine conditions under which this happens.

Theorem 3.6.1 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f_* has an $\mathcal{E}$-indexed right adjoint.*

(ii) *f has a left adjoint when regarded as a morphism $f \to 1_{\mathcal{E}}$ in the 2-category $\mathfrak{Top}/\mathcal{E}$.*

(iii) *f is connected and has a left adjoint in $\mathfrak{Top}$.*

(iv) *f_* has a right adjoint which is full and faithful.*

(v) *f_* has a right adjoint which is a cartesian closed functor.*

(vi) *There exists a geometric morphism $c\colon \mathcal{E} \to \mathcal{F}$ such that $fc \cong 1_{\mathcal{E}}$ and, for every geometric morphism $g\colon \mathcal{G} \to \mathcal{E}$, the composite cg is an initial object of $\mathfrak{Top}/\mathcal{E}\,(g, f)$.*

(vii) *There exists a geometric morphism $c\colon \mathcal{E} \to \mathcal{F}$ such that $fc \cong 1_{\mathcal{E}}$ and there exists a geometric transformation $\alpha\colon cf \to 1_{\mathcal{F}}$ over $\mathcal{E}$.*

Proof (i) $\Rightarrow$ (ii): If c_* is the right adjoint of f_*, then (as observed above) it is the direct image of a geometric morphism, which is left adjoint to f in $\mathfrak{Top}$. But since the adjunction is $\mathcal{E}$-indexed, for any object A of $\mathcal{E}$ we have

$$\begin{aligned} c^*f^*(A) &\cong c^*f^*(\Sigma_A A^*(1)) \\ &\cong \Sigma_A((c^*)^A(f^*)^A A^*(1)) && \text{since } c^* \text{ and } f^* \text{ preserve} \\ & && \mathcal{E}\text{-indexed coproducts} \\ &\cong \Sigma_A A^*(f^*c^*(1)) && \text{since } f^* \text{ and } c^* \text{ are } \mathcal{E}\text{-indexed functors} \\ &\cong \Sigma_A A^*(1) \cong A\,. \end{aligned}$$

So $fc \cong 1_{\mathcal{E}}$, i.e. c is a morphism $1_{\mathcal{E}} \to f$ in $\mathfrak{Top}/\mathcal{E}$; and it is still left adjoint to f in this 2-category, since the unit and counit of the adjunction $(f^* \dashv c^*)$ are $\mathcal{E}$-indexed natural transformations (cf. B3.1.5).

(ii) $\Rightarrow$ (i): Since $\mathfrak{Top}/\mathcal{E}$ is fully embedded in the 2-category $\mathfrak{Top}_{\mathcal{E}}$ of $\mathcal{E}$-indexed toposes by B3.1.5, this is immediate.

(ii) $\Leftrightarrow$ (iii): Condition (ii) says that f_* has a right adjoint and that the composite f_*f^* is isomorphic to the identity functor on $\mathcal{E}$; condition (iii) says additionally that the isomorphism $1_{\mathcal{E}} \cong f_*f^*$ is the unit of the adjunction $(f^* \dashv f_*)$. But, by A1.1.1, this extra condition is automatically satisfied when (ii) holds.

(iii) $\Leftrightarrow$ (iv): Given adjunctions $(f^* \dashv f_* \dashv f^{\#})$, it is well known that f^* is full and faithful iff $f^{\#}$ is, since the composite f_*f^* is left adjoint to $f_*f^{\#}$.

(iv) $\Leftrightarrow$ (v) is immediate from A4.2.9.

(ii) $\Rightarrow$ (vi): Suppose (ii) holds; let $c\colon \mathcal{E} \to \mathcal{F}$ denote the left adjoint of f. Since c is a morphism of $\mathfrak{Top}/\mathcal{E}$, we certainly have $fc \cong 1_{\mathcal{E}}$. Now let $g\colon \mathcal{G} \to \mathcal{E}$ be an $\mathcal{E}$-topos, and $h\colon \mathcal{G} \to \mathcal{F}$ an arbitrary geometric morphism over $\mathcal{E}$. Then 2-cells $cg \to h$ in $\mathfrak{Top}/\mathcal{E}$ correspond to $\mathcal{E}$-indexed natural transformations $g^*c^* \to h^*$.

by B3.1.5, and these in turn correspond to $\mathcal{E}$-indexed natural transformations $f_* \cong c^* \to g_* h^*$. But since f_* is the $\mathcal{E}$-indexed representable functor represented by $1_{\mathcal{F}}$, these in turn correspond to morphisms $1_{\mathcal{E}} \to g_* h^*(1_{\mathcal{F}}) \cong 1_{\mathcal{E}}$, by the $\mathcal{E}$-indexed Yoneda lemma. So there is a unique such 2-cell.

(vi) $\Rightarrow$ (vii) is immediate since (vi) includes the statement that cf is initial in $\mathfrak{Top}/\mathcal{E}\,(\mathcal{F},\mathcal{F})$.

(vii) $\Rightarrow$ (ii): Since α is a 2-cell in $\mathfrak{Top}/\mathcal{E}$, the isomorphism $1_{\mathcal{E}} \to fc$ and α satisfy one of the two triangular identities for an adjunction in $\mathfrak{Top}/\mathcal{E}$; in the terminology of B1.1.8, c is weak left adjoint to f in this 2-category. But idempotent 2-cells split in $\mathfrak{Top}$, by A4.1.15, and an examination of that proof will show that they also split in $\mathfrak{Top}/\mathcal{E}$ for any $\mathcal{E}$; so by B1.1.8 f has a left adjoint in $\mathfrak{Top}/\mathcal{E}$. □

Remark 3.6.2 In the case when f is bounded, we may add a further equivalent condition to 3.6.1, namely

(viii) *f_* preserves arbitrary $\mathcal{E}$-indexed colimits.*

For this condition is always implied by (i), by B1.4.14; and the converse holds if $\mathbb{F}$ has an $\mathcal{E}$-indexed separating family, by B2.4.6.

We say that f is a *local* geometric morphism, or that $\mathcal{F}$ is a *local $\mathcal{E}$-topos*, if it satisfies the equivalent conditions of 3.6.1. (The name could perhaps have been better chosen, given the slightly awkward clash with 'localic morphism'; but both terms are too well established to be easily abandoned.) In the particular case when $\mathcal{E}$ is **Set** (or whatever base topos we happen to be working over at the moment), we simply say $\mathcal{F}$ is a *local topos*. The geometric morphism c of 3.6.1(vi) is often called the *centre* of the local morphism f, for reasons which will become clear shortly.

Examples 3.6.3 (*a*) As we saw in 1.5.6, if X is a locale then $\mathbf{Sh}(X)$ is local (that is, the unique morphism $\mathbf{Sh}(X) \to \mathbf{Set}$ is local) iff X has a focal point, i.e. a point whose only neighbourhood is the whole space. More generally, suppose $f\colon X \to Y$ is a surjective continuous map of (sober) spaces such that, for each point y of Y, the fibre $f^{-1}(y)$ has a focal point $c(y)$, and that the assignment $y \mapsto c(y)$ is continuous as a function $Y \to X$. Then the continuous map c is a left adjoint right inverse for f in the 2-category $\mathfrak{Loc}$ (i.e. $fc = 1_Y$ and $cf \leq 1_X$), and since the assignment $X \mapsto \mathbf{Sh}(X)$ is 2-functorial by 1.4.5, we deduce that the geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ induced by f is local. For example, if L is a normal distributive lattice as defined in [520, II 3.6], and we take X and Y to be the spaces of prime and maximal ideals of L respectively, then the inclusion $Y \to X$ has a (continuous) left inverse $r\colon X \to Y$ sending each prime ideal to the unique maximal ideal which contains it, and the induced geometric morphism $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$ is local.

(*b*) Similarly, if $f\colon \mathcal{C} \to \mathcal{D}$ is any functor between small categories which is a coreflector (i.e. has a left adjoint right inverse), then the induced morphism

$[\mathcal{C}, \mathbf{Set}] \to [\mathcal{D}, \mathbf{Set}]$ is local by A4.1.4. (Indeed, there is a converse result, by A4.1.5: if $\mathcal{C}$ is Cauchy-complete and the morphism induced by f is local, then f is a coreflector, since the left adjoint in $\mathfrak{Top}$ is automatically essential.) In particular, if $\mathcal{C}$ is a small category with an initial object, then the geometric morphism $[\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}$ is local (and the converse holds provided $\mathcal{C}$ is Cauchy-complete).

(*c*) It follows from the preceding sentence that if $\mathbb{T}$ is any cartesian theory, then the classifying topos $\mathbf{Set}[\mathbb{T}]$ is local (over $\mathbf{Set}$); for D3.1.1 shows that $\mathbf{Set}[\mathbb{T}]$ may be identified with $[(\mathcal{C}_{\mathbb{T}})^{\mathrm{op}}, \mathbf{Set}]$, and the syntactic category $\mathcal{C}_{\mathbb{T}}$ has a terminal object by D1.4.2. More generally, for any geometric theory $\mathbb{T}$, a *term model* of $\mathbb{T}$ is defined to be a model M (in $\mathbf{Set}$) such that, for each sort A of the signature of $\mathbb{T}$, each element of MA is the interpretation in M of some closed term of sort A, and two such interpretations s_M and t_M are equal iff the sequent $(\top \vdash_{[]} (s = t))$ is provable in $\mathbb{T}$. It is easy to see that if M is a term model of $\mathbb{T}$, then p^*M is initial in $\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$ for any $\mathbf{Set}$-topos $p\colon \mathcal{E} \to \mathbf{Set}$, and hence by 3.6.1(vi) that the classifying topos $\mathbf{Set}[\mathbb{T}]$ is local. (For a cartesian theory $\mathbb{T}$, the term model is the free model generated by the empty set of generators.)

(*d*) Generalizing the last assertion of (*b*) in a different direction, let $(\mathcal{C}, T)$ be a small site such that $\mathcal{C}$ has a terminal object 1 which is T-irreducible in the sense of 2.2.18(*a*) (i.e. its only T-covering sieve is the maximal sieve M_1). (Such a site is often called a *local site*.) Then we claim that $\mathbf{Sh}(\mathcal{C}, T)$ is a local topos. For the direct image functor $\mathbf{Sh}(\mathcal{C}, T) \to \mathbf{Set}$ is given by evaluation at the terminal object; we know that this has a right adjoint as a functor $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}$, namely right Kan extension along the functor $\mathbf{1} \to \mathcal{C}$ which picks out the terminal object of $\mathcal{C}$; and the irreducibility of the latter ensures that this functor takes values in $\mathbf{Sh}(\mathcal{C}, T)$. (Explicitly, its value at a set A is the functor $U \mapsto A^{\mathcal{C}(1,U)}$, which is a sheaf because $\mathcal{C}(1, -)$ maps covers to epimorphic families; cf. 2.4.10.) Conversely, suppose $\mathcal{E}$ is any bounded local $\mathbf{Set}$-topos, and let $\mathcal{C}$ be a small full subcategory of $\mathcal{E}$ which contains a generating set of objects as well as the terminal object 1. Since the direct image functor $\mathcal{E}(1, -)\colon \mathcal{E} \to \mathbf{Set}$ is also an inverse image functor, it preserves epimorphic families; that is, if $(A_i \to 1 \mid i \in I)$ is an epimorphic family in $\mathcal{E}$, then at least one $A_i \to 1$ must be a split epimorphism. It follows easily that 1 is irreducible for the coverage on $\mathcal{C}$ induced by the canonical coverage on $\mathcal{E}$. Thus we conclude that *a Grothendieck topos is local (over* $\mathbf{Set}$*) iff it can be generated by a local site.* The foregoing arguments are constructive, and so can be interpreted in the internal logic of an arbitrary topos $\mathcal{S}$ to give a characterization of bounded local morphisms with codomain $\mathcal{S}$.

(*e*) If $\mathcal{E}$ is the topos of sheaves on a small full subcategory $\mathcal{C}$ of $\mathbf{Sp}$ as in A2.1.11(*d*), then we saw in 2.3.23 that, for any object X of $\mathcal{C}$, there is a local geometric morphism $\mathcal{E}/l(X) \to \mathbf{Sh}(X)$; and if X is a smooth manifold (resp. an affine scheme of finite type over a field k), then there is a similar local morphism to $\mathbf{Sh}(X)$ from the appropriate slice of the topos of A2.1.11(*e*) (resp. (*f*)).

(*f*) We recall from A2.1.12 that the *Sierpiński cone* on a locally small topos $\mathcal{E}$ is the topos obtained by glueing along the functor $\mathcal{E}(1, -)\colon \mathcal{E} \to \mathbf{Set}$. More generally, given a geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$, we define the Sierpiński cone on $\mathcal{E}$

over $\mathcal{S}$ to be the topos obtained by glueing along $p_*: \mathcal{E} \to \mathcal{S}$; we shall denote it by $\mathbf{scn}_{\mathcal{S}}(\mathcal{E})$. By B3.4.2, it may be viewed as the cocomma object $(p \uparrow 1_{\mathcal{E}})$ in $\mathfrak{Top}$; that is, geometric morphisms $\mathbf{scn}_{\mathcal{S}}(\mathcal{E}) \to \mathcal{F}$ correspond to triples (f, g, α) where $f: \mathcal{E} \to \mathcal{F}$, $g: \mathcal{S} \to \mathcal{F}$ and $\alpha: pg \to f$ is a geometric transformation. In particular we have a morphism $q: \mathbf{scn}_{\mathcal{S}}(\mathcal{E}) \to \mathcal{S}$ corresponding to the triple $(p, 1_{\mathcal{S}}, 1_p)$; we claim that this is always local. For its inverse image is easily seen to be given by $q^*(I) = (p^*I, I, \eta_I)$ where η is the unit of $(p^* \dashv p_*)$; hence morphisms $q^*(I) \to (A, J, f)$ are pairs (u, v) where $u: p^*(I) \to A$ and $v: I \to J$ make the square

$$\begin{tikzcd}
I \arrow[r, "\eta_I"] \arrow[d, "v"'] & p_*p^*(I) \arrow[d, "p_*(u)"] \\
J \arrow[r, "f"] & p_*(A)
\end{tikzcd}$$

commute. But this is equivalent to saying that the transpose of u is the composite fv, so the pair (u, v) is uniquely determined by v, and we see that the right adjoint q_* of q^* coincides with the projection $\mathbf{scn}_{\mathcal{S}}(\mathcal{E}) \to \mathcal{S}$, i.e. the inverse image of the closed inclusion $c: \mathcal{S} \to \mathbf{scn}_{\mathcal{S}}(\mathcal{E})$ (cf. A4.5.6). We note also that the composite of q with the open inclusion $d: \mathcal{E} \to \mathbf{scn}_{\mathcal{S}}(\mathcal{E})$ is our original morphism p; thus we have shown that *every geometric morphism may be factored as an open inclusion followed by a local morphism.*

Following on from the last example above, we note

Lemma 3.6.4 *For an $\mathcal{S}$-topos $p: \mathcal{E} \to \mathcal{S}$, the following are equivalent:*

(i) *p is local.*

(ii) *$\mathcal{E}$ is a retract of a local $\mathcal{S}$-topos.*

(iii) *$\mathcal{E}$ is a retract of $\mathbf{scn}_{\mathcal{S}}(\mathcal{F})$ for some $\mathcal{S}$-topos $\mathcal{F}$.*

(iv) *$\mathcal{E}$ is an adjoint retract (cf.* B1.1.9(c)*) of $\mathbf{scn}_{\mathcal{S}}(\mathcal{F})$ for some $\mathcal{S}$-topos $\mathcal{F}$.*

(v) *The open inclusion $d: \mathcal{E} \to \mathbf{scn}_{\mathcal{S}}(\mathcal{E})$ has a left adjoint.*

Proof (i) $\Rightarrow$ (v): Suppose $\mathcal{E}$ is a local $\mathcal{S}$-topos with centre c. Then the morphisms $1_{\mathcal{E}}$ and c, together with the geometric transformation $cp \to 1_{\mathcal{E}}$ of 3.6.1(vi), induce a morphism $\mathbf{scn}_{\mathcal{S}}(\mathcal{E}) \to \mathcal{E}$ over $\mathcal{S}$, which is a splitting for the open inclusion. Moreover, it is left adjoint to d: since $c^* \cong p_*$, its inverse image may be identified with the functor $A \mapsto (A, p_*A, 1_{p_*A})$, which is the direct image of d.

(v) $\Rightarrow$ (iv) is clear, since any left adjoint for an inclusion is necessarily a one-sided inverse for it (up to isomorphism). (iv) $\Rightarrow$ (iii) is trivial; and (iii) $\Rightarrow$ (ii) is immediate from 3.6.3(f). Finally, (ii) $\Rightarrow$ (i) follows easily from B1.1.10 (and the splittability of idempotent 2-cells in $\mathfrak{Top}/\mathcal{S}$, cf. A4.1.15). $\square$

Remark 3.6.5 By applying the argument in the proof of 3.6.4(i)$\Rightarrow$(v) to local $\mathcal{S}$-toposes of the form $\mathbf{scn}_{\mathcal{S}}(\mathcal{F})$, we obtain the multiplication of a KZ-monad structure on the functor $\mathbf{scn}_{\mathcal{S}}(-)\colon \mathfrak{Top}/\mathcal{S} \to \mathfrak{Top}/\mathcal{S}$, whose unit is the open inclusion. Lemma 3.6.4, in conjunction with B1.1.14, now tells us that the algebras for this monad are exactly the local $\mathcal{S}$-toposes (and morphisms of algebras are morphisms between local $\mathcal{S}$-toposes which commute up to isomorphism with their centres).

If the ideas in the preceding paragraph seem reminiscent of the notion of hyperlocal topos which we studied in Section B4.4, this is no accident. Of course, it follows immediately from B4.4.19(vi) and 3.6.1(ii) that hyperlocal $\mathcal{S}$-toposes are local. But it is also the case that $\mathbf{scn}_{\mathcal{S}}(-)$ may, like the lower bagdomain functor B_L, be regarded as a partial product functor. We note that, for any $\mathcal{S}$-topos $\mathcal{F}$, a morphism $\mathcal{F} \to \mathbf{scn}_{\mathcal{S}}(\mathcal{E})$ over $\mathcal{S}$ corresponds, via pullback along the open inclusion d, to a pair (U, v) where U is a subterminal object of $\mathcal{F}$ and v is a morphism $\mathcal{F}/U \to \mathcal{E}$ over $\mathcal{S}$; in other words, it is just a 'bag of $\mathbb{T}$-models', where $\mathbb{T}$ is the geometric theory classified by $\mathcal{E}$, whose indexing object is subterminal. (Equivalently again, it corresponds to a pre-geometric morphism (A4.1.13) $\mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ whose 'inverse image' functor preserves binary products as well as pullbacks.) From this description, it is easy to see that we have $\mathbf{scn}_{\mathcal{S}}(-) \simeq \mathcal{P}_{\bullet}(-, d_1)$, where $d_1\colon \mathcal{S} \to [\mathbf{2}, \mathcal{S}] \simeq \mathbf{scn}_{\mathcal{S}}(\mathcal{S})$ is the open point of the Sierpiński topos over $\mathcal{S}$ (the classifying topos for the theory of subterminal objects, cf. B3.2.11).

From another of the examples in 3.6.3, we obtain

Scholium 3.6.6 *A geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ between bounded $\mathcal{S}$-toposes is local iff it can be induced by a fibration of sites $(P, T)\colon (\mathbb{D}, K) \to (\mathbb{C}, J)$ in $\mathcal{S}$ such that T reflects (as well as preserves) covers.*

Proof If $\mathcal{F}$ is local, we may choose a local site $(\mathbb{B}, T)$ for it in $\mathcal{E}$, by 3.6.3(d). If we then express $\mathcal{E}$ as $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J)$ and externalize the site $(\mathbb{B}, T)$ to obtain a fibration of toposes as in 2.5.7, the condition that the terminal object of $\mathbb{B}$ is T-irreducible easily translates into the assertion that T reflects covers.

Conversely, if f is induced by a fibration of sites as in the statement, then since T is both cover-reflecting and a morphism of sites it induces an adjoint pair of geometric morphisms as in 2.3.23; also, since T is full and faithful, the left adjoint $c\colon \mathcal{E} \to \mathcal{F}$ is an inclusion (equivalently, f is connected). So f is local by 3.6.1(iii). □

Next, we investigate the stability properties of local maps.

Lemma 3.6.7

(i) *Local maps are tidy.*

(ii) *A composite of local maps is local.*

(iii) *In a commutative triangle*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

of geometric morphisms, if h is local and g is connected, then f is local. Moreover, if c is the centre of h, then gc is the centre of f.

(iv) *Let*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

be a pullback square in $\mathfrak{Top}$ with f local. Then k is local. Moreover, if c is the centre of f, then the centre of k is the morphism into the pullback induced by cg and $1_{\mathcal{G}}$.

(v) *Let*

$$\cdots \xrightarrow{f_3} \mathcal{E}_2 \xrightarrow{f_2} \mathcal{E}_1 \xrightarrow{f_1} \mathcal{E}_0$$

*be an inverse sequence of toposes and local morphisms having a (**Cat**-enriched) limit in $\mathfrak{Top}$ (or in $\mathfrak{BTop}/\mathcal{S}$ for some $\mathcal{S}$). Then the legs $f^n_\infty : \mathcal{E}_\infty \to \mathcal{E}_n$ of the limit cone are local.*

Proof (i) is immediate from 3.6.2, and (ii) is immediate from either (iii) or (iv) of 3.6.1.

(iii): Since the unit of $(g^* \dashv g_*)$ is an isomorphism, we have $f_*(B) \cong f_*g_*g^*(B) \cong h_*g^*(B) \cong c^*g^*(B)$ for any object B of $\mathcal{F}$. So f_* is indeed an inverse image functor; and f is connected, since we have $f_*f^* \cong c^*g^*f^* \cong c^*h^* \cong 1_{\mathcal{E}}$.

(iv): For this we use 3.6.1(vi). Let $d: \mathcal{G} \to \mathcal{H}$ be the morphism induced by cg and $1_{\mathcal{G}}$. Then, for any $l: \mathcal{L} \to \mathcal{G}$, composition with h induces an equivalence between $\mathfrak{Top}/\mathcal{G}\,(l,k)$ and $\mathfrak{Top}/\mathcal{E}\,(gl,f)$; so from the fact that $cgl \cong hdl$ is initial in the latter we deduce that dl is initial in the former.

(v) Let $c_n: \mathcal{E}_{n-1} \to \mathcal{E}_n$ denote the centre of f_n, and, for $n < m$, let $c^n_m: \mathcal{E}_n \to \mathcal{E}_m$ denote the composite of the c_j with $n+1 \leq j \leq m$. Since the f_j are all connected, it is easy to verify that for each n we have a (pseudo-)cone over the diagram with vertex $\mathcal{E}_n$, whose mth component is c^n_m if $m > n$, and f^m_n (the corresponding composite of the f_j) if $m \leq n$. These induce a morphism $c^n_\infty: \mathcal{E}_n \to \mathcal{E}_\infty$; it is clear that $f^n_\infty c^n_\infty$ is isomorphic to the identity, and

using the fact that $\mathcal{E}_\infty$ is a **Cat**-enriched limit it is easy to construct a geometric transformation from $c_\infty^n f_\infty^n$ to the identity, which is the counit of an adjunction $(c_\infty^n \dashv f_\infty^n)$. □

Remarks 3.6.8 (*a*) The hypothesis 'g is connected' in 3.6.7(iii) cannot be weakened to 'g is surjective', even if we add the further hypothesis that g is open (or even atomic). Let $\mathcal{C}$ be the finite category with three objects U, V, W and two non-identity morphisms $U \to V \leftarrow W$, and let $\mathcal{D}$ be the quotient of $\mathcal{C}$ obtained by identifying the objects U and W (but not the two non-identity morphisms). The quotient functor $\mathcal{C} \to \mathcal{D}$ is surjective on objects and a discrete fibration, so the induced geometric morphism $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}] \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is a surjective local homeomorphism. However, $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is local over **Set**, since $\mathcal{C}$ has a terminal object; but $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is not local, since $\mathcal{D}$ is Cauchy-complete but does not have a terminal object. (In fact the direct image functor $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}] \to \mathbf{Set}$ sends a diagram of shape $\mathcal{D}^{\mathrm{op}}$ (that is, a parallel pair) to its equalizer, and it is easily seen that this functor fails to preserve epimorphisms.)

(*b*) Similarly, 3.6.7(iii) does not have a 'dual': even if we assume that both f and h are local in the diagram of 3.6.7(iii), and further that g has a right inverse, there is still no reason why g should be local. To see this, we use 3.6.3(*c*): we take $\mathcal{E}$ to be **Set**, and $\mathcal{F}$ and $\mathcal{G}$ to be the classifying toposes for the theories of abelian groups and of rings (not necessarily with 1), respectively. Then both $\mathcal{F}$ and $\mathcal{G}$ are local over $\mathcal{E}$, and we have a morphism $g \colon \mathcal{G} \to \mathcal{F}$ classifying the underlying additive group of the generic ring, which has a right inverse obtained by equipping the generic abelian group with the zero multiplication. However, g is not connected, and therefore not local; it does have a left adjoint in $\mathfrak{Top}$, corresponding to the free ring generated by the generic abelian group, but the unit of the adjunction is not an isomorphism.

(*c*) It will be noted that, for once, we have been able to prove the pullback-stability theorem (3.6.7(iv)) without worrying about which of the maps in the pullback is bounded. However, if in the situation of 3.6.7(iv) we happen to know that f is bounded as well as local, we can give an alternative proof of the result using the site characterization of bounded morphisms given in 3.6.3(*d*), since (as we observed in the proof of 3.2.21) the assertion that a particular object of a site is irreducible is stable under change of base. Note also that the assertion in 3.6.7(iv) that the centre of a local morphism is 'stable under pullback' easily yields the fact that local morphisms are SRBC morphisms in the sense of 2.4.16 (which we knew in any case, since local morphisms are tidy).

(*d*) Similarly, in proving the stability of local maps under inverse limits, we did not need to use our characterization of local maps in terms of fibrations of sites (3.6.6), although it could no doubt be derived using the latter. We note also that the argument of 3.6.7(v) works for limits over more general cofiltered categories, although we shall also be able to derive the latter fact by way of the descent theorem for open surjections (see 5.1.13 below), as we do for other properties of geometric morphisms.

Next, we note a useful 'orthogonality property' of local morphisms. We say an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ (or the geometric morphism p) is *grouplike* if, for any $\mathcal{S}$-topos $\mathcal{F}$, the category $\mathfrak{Top}/\mathcal{S}\,(\mathcal{F},\mathcal{E})$ is a groupoid. If $\mathcal{E}$ is bounded over $\mathcal{S}$, this says that it classifies a geometric theory $\mathbb{T}$ such that every homomorphism of $\mathbb{T}$-models (in arbitrary $\mathcal{S}$-toposes) is an isomorphism.

Examples 3.6.9 (*a*) Let G be an internal group in a topos $\mathcal{S}$. In B3.2.4(*b*) we gave a description of the notion of G-torsor, from which it follows that $[G,\mathcal{S}]$ is a grouplike $\mathcal{S}$-topos. The same is true of $[\mathbb{G},\mathcal{S}]$ for any internal groupoid $\mathbb{G}$ in $\mathcal{S}$. (However, the result does not extend to $\mathbf{Cont}(G)$ for a topological group G: for example, if G is the group of permutations of $\mathbb{N}$, with the topology of pointwise convergence, then $\mathbf{Cont}(G)$ classifies the theory of infinite decidable objects (D3.4.10), and so any injection $\mathbb{N} \to \mathbb{N}$ which is not surjective defines a non-invertible endomorphism of a point of $\mathbf{Cont}(G)$.)

(*b*) If X is a T_U-locale in a topos $\mathcal{S}$ (cf. 1.2.17), then $\mathbf{Sh}_{\mathcal{S}}(X)$ is a grouplike $\mathcal{S}$-topos, since geometric morphisms $\mathcal{E} \to \mathbf{Sh}_{\mathcal{S}}(X)$ over $\mathcal{S}$ correspond to locale maps $Y \to X$, where Y is (the locale in $\mathcal{S}$ corresponding to) the localic reflection of $\mathcal{E}$.

Proposition 3.6.10 *Local morphisms are orthogonal (in the 2-categorical sense) to grouplike morphisms: that is, if $p\colon \mathcal{E} \to \mathcal{S}$ is a grouplike $\mathcal{S}$-topos and $f\colon \mathcal{G} \to \mathcal{F}$ is a local morphism of $\mathcal{S}$-toposes, then the functor*

$$\mathfrak{Top}/\mathcal{S}\,(\mathcal{F},\mathcal{E}) \longrightarrow \mathfrak{Top}/\mathcal{S}\,(\mathcal{G},\mathcal{E})$$

induced by composition with f is part of an equivalence of categories.

Proof The displayed functor has a right adjoint, induced by composition with the centre of f. But any adjunction between groupoids is an equivalence. □

As an example of the applications of 3.6.10, we give

Corollary 3.6.11 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be a local geometric morphism over some base topos $\mathcal{S}$ with a natural number object. Then f_* preserves the object R_d of Dedekind real numbers (cf.* D4.7.5*).*

Proof Let A be an object of $\mathcal{E}$. By definition, morphisms $A \to R_d$ in $\mathcal{E}$ correspond to Dedekind real numbers in $\mathcal{E}/A$, and hence to geometric morphisms $\mathcal{E}/A \to \mathbf{Sh}_{\mathcal{S}}(R_f)$ over $\mathcal{S}$, where R_f is the locale of 'formal real numbers' in $\mathcal{S}$ (cf. D4.7.4). But R_f is regular and hence T_U by 1.2.17, and $\mathcal{F}/f^*A \to \mathcal{E}/A$ is local by 3.6.7(iv), so these correspond to morphisms $\mathcal{F}/f^*A \to \mathbf{Sh}_{\mathcal{S}}(R_f)$ over $\mathcal{S}$, and hence to morphisms $f^*A \to R_d$ in $\mathcal{F}$. Thus f_*R_d represents the same functor as the Dedekind real number object in $\mathcal{E}$. □

We turn now to the notion of localization. Given a category $\mathcal{C}$ and an object U of $\mathcal{C}$, we have a universal way of 'making U into an initial object of $\mathcal{C}$': namely, forming the co-slice category $U\backslash\mathcal{C}$. (Formally, the forgetful functor $U\backslash\mathcal{C} \to \mathcal{C}$ is

universal amongst functors from categories with initial objects to $\mathcal{C}$, which send the initial object to U; cf. A1.2.9.) In the same vein, given a space X and a point x of X, there is a largest subspace of X which contains x as a focal point, namely the subspace of all points y such that $x \leq y$ in the specialization ordering. These observations inspire the proof of the following result:

Theorem 3.6.12 *Given a bounded $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ and a particular section $s\colon \mathcal{S} \to \mathcal{E}$ of p, there exists a bounded local $\mathcal{S}$-topos $p_s\colon \mathcal{E}_s \to \mathcal{S}$ (called the* localization *of $\mathcal{E}$ at s) equipped with a geometric morphism $\mathcal{E}_s \to \mathcal{E}$ over S which carries the centre of $\mathcal{E}_s$ into the chosen section s, and which induces an equivalence*

$$\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E}_s) \simeq sq\backslash(\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E}))$$

for any bounded $\mathcal{S}$-topos $(q\colon \mathcal{F} \to \mathcal{S})$.

Proof We need to recall a result from Part B: namely, the fact that the 2-category of bounded $\mathcal{S}$-toposes admits cotensors with $\mathbf{2}$ (B4.1.4). That is, for any bounded $\mathcal{S}$-topos $\mathcal{E}$, there exists a topos $\mathcal{G} = \mathbf{2} \pitchfork \mathcal{E}$ equipped with a diagram

$$\begin{tikzcd} \mathcal{G} \arrow[r, shift left=1.5ex, "d_1"] \arrow[r, shift right=1.5ex, "d_0"'] \arrow[r, phantom, "\downarrow"] & \mathcal{E} \end{tikzcd}$$

which is universal amongst 2-cells with codomain $\mathcal{E}$. Now the identity 2-cell on the identity geometric morphism $\mathcal{E} \to \mathcal{E}$ induces a geometric morphism $\Delta\colon \mathcal{E} \to \mathcal{G}$, which clearly satisfies $d_0\Delta \cong d_1\Delta \cong 1_{\mathcal{E}}$; and there are canonical geometric transformations $\Delta d_1 \to 1_{\mathcal{G}} \to \Delta d_0$, which make Δ left adjoint to d_1 and right adjoint to d_0 in $\mathfrak{Top}/\mathcal{S}$. In particular, d_1 is a local morphism, with centre Δ. Thus if we form the pullback

$$\begin{tikzcd} \mathcal{E}_s \arrow[r, "\tilde{s}"] \arrow[d, "p_s"] & \mathbf{2} \pitchfork \mathcal{E} \arrow[d, "d_1"] \\ \mathcal{S} \arrow[r, "s"] & \mathcal{E} \end{tikzcd}$$

then $\mathcal{E}_s$ is local over $\mathcal{S}$ by 3.6.7(iv). Moreover, for any bounded $\mathcal{S}$-topos $q\colon \mathcal{F} \to \mathcal{S}$, morphisms $f\colon \mathcal{G} \to \mathcal{E}_s$ over $\mathcal{S}$ correspond to morphisms $g = \tilde{s}f\colon \mathcal{F} \to \mathbf{2} \pitchfork \mathcal{E}$ satisfying $d_1 g = sq$, and hence to morphisms $h = d_0 g\colon \mathcal{F} \to \mathcal{E}$ equipped with a geometric transformation $sq \to h$ over $\mathcal{S}$. □

There are two alternative ways of representing $\mathcal{E}_s$ as the solution of a universal problem. For the first, let $\mathbf{p}\mathfrak{Top}/\mathcal{S}$ denote the 2-category whose objects are bounded $\mathcal{S}$-toposes $(p\colon \mathcal{E} \to \mathcal{S})$ equipped with a specified point s (that is, a

section of p), and whose 1-cells are geometric morphisms over $\mathcal{S}$ preserving the specified points; 2-cells between morphisms $f, f' : (p, s) \rightrightarrows (p', s')$ are geometric transformations $\alpha : f \to f'$ such that both $p' \circ \alpha$ and $\alpha \circ s$ are (the canonical) isomorphisms.

Corollary 3.6.13 *The assignment $(p : \mathcal{E} \to \mathcal{S}, s) \mapsto \mathcal{E}_s$ is a functor $\mathfrak{pTop}/\mathcal{S} \to \mathfrak{BTop}/\mathcal{S}$, right adjoint to the functor which sends $p : \mathcal{E} \to \mathcal{S}$ to $\mathbf{scn}_{\mathcal{S}}(\mathcal{E})$ equipped with its closed point.*

Proof For an $\mathcal{S}$-topos $(q : \mathcal{F} \to \mathcal{S})$, morphisms $\mathcal{F} \to \mathcal{E}_s$ over $\mathcal{S}$ correspond to morphisms $f : \mathcal{F} \to \mathbf{2} \pitchfork \mathcal{E}$ such that $d_1 f \cong sq$, and hence to pairs (g, α) where $g = d_0 f$ is a morphism $\mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$, and α is a geometric transformation $sq \to g$. But these in turn correspond to geometric morphisms $\mathbf{scn}_{\mathcal{S}}(\mathcal{F}) \to \mathcal{E}$ over $\mathcal{S}$ carrying the closed point of $\mathbf{scn}_{\mathcal{S}}(\mathcal{F})$ to s. □

Another, more 'relaxed', way of making pointed bounded $\mathcal{S}$-toposes into a 2-category is to allow the 1-cells $(p, s) \to (p', s')$ to be pairs (f, α) where f is a morphism over $\mathcal{S}$ (that is, $p'f \cong p$) but $\alpha : s' \to fs$ is a geometric transformation over $\mathcal{S}$ which is not assumed to be invertible. With the appropriate relaxation of the notion of 2-cell, this yields a 2-category which we shall denote by $\mathfrak{PTop}/\mathcal{S}$. We write $\mathfrak{LocTop}/\mathcal{S}$ for the full sub-2-category of $\mathfrak{BTop}/\mathcal{S}$ whose objects are local $\mathcal{S}$-toposes; by 3.6.1(vi) we see that the operation of equipping a local topos with its centre defines a full embedding of 2-categories $\mathfrak{LocTop}/\mathcal{S} \to \mathfrak{PTop}/\mathcal{S}$.

Corollary 3.6.14 *The assignment which sends $(p : \mathcal{E} \to \mathcal{S}, s)$ to $\mathcal{E}_s$ is a functor $\mathfrak{PTop}/\mathcal{S} \to \mathfrak{LocTop}/\mathcal{S}$, right adjoint to the full embedding just defined.*

Proof First we indicate how $(\mathcal{E}, s) \mapsto \mathcal{E}_s$ becomes a functor on $\mathfrak{PTop}/\mathcal{S}$. Let $(f, \alpha) : (\mathcal{F}, t) \to (\mathcal{E}, s)$ be a 1-cell in this 2-category; then we have a diagram

$$\begin{tikzcd}
\mathcal{F}_t \arrow[r] \arrow[dd] & \mathbf{2} \pitchfork \mathcal{F} \arrow[d, "d_1"'] \arrow[dr, "d_0"] \arrow[dr, phantom, shift right=3ex, "\rightarrow"] & \\
& \mathcal{F} \arrow[r, "1"'] \arrow[dr, phantom, "\uparrow \alpha"] & \mathcal{F} \arrow[d, "f"] \\
\mathcal{S} \arrow[ur, "t"] \arrow[rr, "s"] & & \mathcal{E}
\end{tikzcd}$$

which on pasting yields a diagram of the required form to induce a morphism $\mathcal{F}_t \to \mathcal{E}_s$ over $\mathcal{S}$. The definition of the localization functor on 2-cells of $\mathfrak{PTop}/\mathcal{S}$ is straightforward.

To establish the adjunction, let $(q : \mathcal{F} \to \mathcal{S})$ be a local $\mathcal{S}$-topos, with centre c, and suppose we are given $(f, \alpha) : (\mathcal{F}, c) \to (\mathcal{E}, s)$ in $\mathfrak{PTop}/\mathcal{S}$. Then we may

obtain a morphism $\mathcal{F} \to \mathcal{E}_s$ from the diagram

$$\begin{tikzcd}
\mathcal{F} \arrow[r, "1"] \arrow[dr, "q"'] \arrow[dr, phantom, "\uparrow", near start, shift left=4ex] & \mathcal{F} \arrow[r, "f"] \arrow[dr, phantom, "\uparrow\alpha", near start] & \mathcal{E} \\
& \mathcal{S} \arrow[u, "c"'] \arrow[ur, "s"'] &
\end{tikzcd} .$$

Conversely, given $g\colon \mathcal{F} \to \mathcal{E}_s$ over $\mathcal{S}$, the diagram

$$\begin{tikzcd}
\mathcal{F} \arrow[r, "g"] \arrow[dr, "q"] & \mathcal{E}_s \arrow[r] \arrow[d] & \mathbf{2} \pitchfork \mathcal{E} \arrow[r, "d_0"] \arrow[d, "d_1"'] \arrow[dr, phantom, "\uparrow"] & \mathcal{E} \\
\mathcal{S} \arrow[u, "c"] \arrow[r, "1"] & \mathcal{S} \arrow[r, "s"] & \mathcal{E} \arrow[ur, "1"'] &
\end{tikzcd}$$

yields a morphism $(\mathcal{F}, c) \to (\mathcal{E}, s)$ in $\mathfrak{PTop}/\mathcal{S}$. The rest of the proof is straightforward verification. □

Examples 3.6.15 (*a*) If $p\colon \mathcal{E} \to \mathcal{S}$ is a grouplike $\mathcal{S}$-topos, then the domain map $\mathbf{2} \pitchfork \mathcal{E} \to \mathcal{E}$ is an equivalence – indeed, this property characterizes grouplike $\mathcal{S}$-toposes, since a category $\mathcal{C}$ is a groupoid iff the domain map $[\mathbf{2}, \mathcal{C}] \to \mathcal{C}$ is an equivalence. It is therefore immediate from the definition that any localization of a grouplike $\mathcal{S}$-topos is equivalent to the base topos $\mathcal{S}$.

(*b*) If $p\colon \mathcal{E} \to \mathcal{S}$ is already local, and c is its centre, then the localization $\mathcal{E}_c$ is simply (equivalent to) $\mathcal{E}$ itself, by 3.6.14.

(*c*) If $\mathcal{C}$ is a small category and U is an object of $\mathcal{C}$, then the localization of $[\mathcal{C}, \mathbf{Set}]$ at the point induced (as in A4.1.4) by the functor $U\colon \mathbf{1} \to \mathcal{C}$ is simply the functor category $[U\backslash\mathcal{C}, \mathbf{Set}]$; for it is easy to verify that the notion of a $\mathcal{C}$-torsor F equipped with a natural transformation from the functor $\mathcal{C}(-, U)$ to F (corresponding to an element $x \in F(U)$) is equivalent to that of a torsor G on $U\backslash\mathcal{C}$, by an argument like that in the proof of B4.1.2.

(*d*) If X is the spectrum of a commutative ring A in $\mathbf{Set}$ (i.e. the space of prime ideals of A, equipped with the topology generated by the open sets $D(a) = \{P \in X \mid a \notin P\}$, $a \in A$ – it is well known that this topology is isomorphic to the frame of radical ideals of A) and P is a particular prime ideal, then $\mathbf{Sh}(X)_P$ may be identified with the spectrum of the ring A_P (the *localization* of A at P) obtained by formally adjoining inverses for all elements of $A \setminus P$. Of course, the topos-theoretic notion of localization was originally inspired by this algebraic example.

(*e*) Finally, we mention an example from [545] of one way to define the notion of 'the germ of a manifold at a point' in topos-theoretic terms. Let M be a smooth

manifold, and x a point of M; if we simply localize the topos $\mathbf{Sh}(M)$ at the point x, we will obtain the topos **Set** by (a) above, since M is a Hausdorff space and so $\mathbf{Sh}(M)$ is grouplike. However, we have another topos associated with M, namely the slice $\mathcal{E}/l(M)$, where $\mathcal{E}$ is the topos constructed as in A2.1.11(e) from a small full subcategory $\mathcal{C}$ of **Mf** containing M, and $l\colon \mathcal{C} \to \mathcal{E}$ is the canonical functor. As we saw in 3.6.3(e) above, there is a local morphism $\mathcal{E}/l(M) \to \mathbf{Sh}(M)$; if c denotes the centre of this morphism, then the point cx of $\mathcal{E}/M$ corresponds to the $(\mathcal{C}/M)^{\mathrm{op}}$-torsor whose value at $f\colon N \to M$ is simply the set $f^{-1}(x)$. Localizing at this point is clearly unlikely to tell us anything about the topological structure of M at x; and indeed, it is easy to show that $(\mathcal{E}/l(M))_{cx} \simeq \mathbf{Sh}(M)_x \simeq \mathbf{Set}$. However, we have another point $\tilde{x}$ of $\mathcal{E}/l(M)$, corresponding to the torsor which sends $f\colon N \to M$ to the set of germs at x of continuous sections of f. Clearly, evaluation of germs at x yields a natural transformation from this functor to the one considered previously, and hence to a geometric transformation $\tilde{x} \to cx$ (so that cx will 'appear' as a point of the localization $(\mathcal{E}/l(M))_{\tilde{x}}$). In [545] it is shown, by constructing an explicit site of definition for this localization, that it does indeed represent 'the germ of M at x', at least in the sense that, for any object N of $\mathcal{C}$, the category $\mathfrak{BTop}/\mathcal{E}\,((\mathcal{E}/l(M))_{\tilde{x}}, \mathcal{E}/l(N))$ is equivalent to the (discrete) set of germs at x of smooth maps $M \to N$.

As promised, we conclude this section by briefly reviewing the duals of local maps. Paralleling 3.6.1, we have

Theorem 3.6.16 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is locally connected and the left adjoint $f_!$ of f^* is cartesian.*

(ii) *f has a right adjoint when regarded as a morphism $f \to 1_{\mathcal{E}}$ in $\mathfrak{Top}/\mathcal{E}$.*

(iii) *f is connected and has a right adjoint in $\mathfrak{Top}$.*

(iv) *There exists a geometric morphism $d\colon \mathcal{E} \to \mathcal{F}$ such that $fd \cong 1_{\mathcal{E}}$ and, for every $\mathcal{E}$-topos $g\colon \mathcal{G} \to \mathcal{E}$, the composite dg is a terminal object of $\mathfrak{Top}/\mathcal{E}\,(g, f)$.*

(v) *f has a section d for which there exists a geometric transformation $\beta\colon 1_{\mathcal{F}} \to df$ over $\mathcal{E}$.*

Proof (i) $\Rightarrow$ (iii): If (i) holds, then f is connected by 3.3.3; and $f_!$ is the inverse image of a geometric morphism, which is right adjoint to f in $\mathfrak{Top}$.

(iii) $\Rightarrow$ (ii) is similar to the implication (iii) $\Rightarrow$ (ii) of 3.6.1.

(ii) $\Rightarrow$ (i): If f has a right adjoint d in $\mathfrak{Top}/\mathcal{E}$, then d^* is an $\mathcal{E}$-indexed left adjoint for f^*. So f is locally connected, and (i) holds.

(ii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (ii) is exactly dual to the proof of (ii) $\Rightarrow$ (vi) $\Rightarrow$ (vii) $\Rightarrow$ (ii) in 3.6.1. □

We say that a geometric morphism is *totally connected* if it satisfies the equivalent conditions of 3.6.16. Note that the morphism d of 3.6.16(iv) is necessarily a

dense inclusion (indeed, a flat inclusion in the sense of 1.1.16(d)), since $d_* = f^*$ is full and faithful and preserves finite unions.

Examples 3.6.17 (a) For a space X, we saw that $\mathbf{Sh}(X)$ is local over **Set** iff X has a focal point, i.e. one whose only neighbourhood is the whole space. Similarly, $\mathbf{Sh}(X)$ is totally connected iff X has a dense point d – one whose closure is the whole space. This condition implies that every inhabited open set is connected (which explains the use of the term 'totally connected'); and the converse holds provided X is sober, since the condition 'every inhabited open set is connected' implies that the inhabited open sets form a completely prime filter in $\mathcal{O}(X)$, and the corresponding point of X is clearly dense. To see that this implies that $\mathbf{Sh}(X)$ is a totally connected topos, observe that it clearly implies that X is locally connected, and hence so is the domain of any local homeomorphism $E \to X$; now each connected component of E must meet the fibre over the dense point d in exactly one point, and so we may identify d^* with the 'connected components' functor $\gamma_!$ of 1.5.9.

(b) Paralleling 3.6.3(b), we see that if a small category $\mathcal{C}$ has a terminal object, then $[\mathcal{C}, \mathbf{Set}]$ is a totally connected **Set**-topos. However, this condition is not necessary, since the right adjoint d need not be induced by a functor $\mathbf{1} \to \mathcal{C}$ (that is, an object of $\mathcal{C}$). In fact, as we know from B2.6.8, a diagram category $[\mathbb{C}, \mathcal{S}]$ is totally connected over $\mathcal{S}$ iff $\mathbb{C}$ is filtered.

(c) Generalizing the last example, let $(\mathbb{C}, J)$ be an internal site in a topos $\mathcal{S}$, such that $\mathbb{C}^{\mathrm{op}}$ is filtered (for example, it would be sufficient for $\mathbb{C}$ to be cartesian) and every J-covering family is connected. Then by 3.3.10 we know that constant diagrams of shape $\mathbb{C}^{\mathrm{op}}$ are sheaves, and hence the left adjoint $p_!$ of the inverse image functor $p^* : \mathcal{S} \to \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J)$ is simply $\varinjlim_{\mathbb{C}^{\mathrm{op}}}$ applied to sheaves on $\mathbb{C}$: in particular, it is cartesian by B2.6.8 and the fact that the inclusion $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, J) \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ preserves finite limits. So $\mathbf{Sh}_{\mathcal{S}}(\mathcal{C}, J)$ is totally connected over $\mathcal{S}$. Conversely, if $\mathcal{S}$ has a natural number object, then every totally connected bounded $\mathcal{S}$-topos $\mathcal{E}$ can be generated by a site of this form, since the connected objects of $\mathcal{E}$ (i.e. those objects A satisfying $p_!(A) \cong 1$) are closed under finite limits.

(d) We may also dualize 3.6.3(f): if we glue along the inverse image $p^* : \mathcal{S} \to \mathcal{E}$ of a geometric morphism p, instead of the direct image – that is, if we form the cocomma object $(1_{\mathcal{E}} \uparrow p)$ in $\mathfrak{Top}$, instead of $(p \uparrow 1_{\mathcal{E}})$ – then we obtain a topos which is readily seen to be totally connected over $\mathcal{S}$. In particular, *any geometric morphism can be factored as a closed inclusion followed by a totally connected morphism.* (It will be noted that we have already exploited this factorization in the proof of 3.3.14.) And, dualizing 3.6.4, we see that p itself is totally connected iff $\mathcal{E}$ is a retract of $(1_{\mathcal{E}} \uparrow p)$ in $\mathfrak{Top}/\mathcal{S}$, iff it is a coadjoint retract of $(1_{\mathcal{E}} \uparrow p)$. Note also that the functor $(p : \mathcal{E} \to \mathcal{S}) \mapsto (1_{\mathcal{E}} \uparrow p)$ may be regarded as a contravariant partial product functor in the sense of B4.4.13, using the closed inclusion $\mathcal{S} \to [\mathbf{2}, \mathcal{S}]$ in place of the open inclusion which we used in 3.6.5. In particular, since closed inclusions are entire, one may deduce that *any bounded $\mathcal{S}$-topos which*

admits an algebra structure for the upper bagdomain monad of B4.4.21 *is totally connected.*

(e) Our final example has no parallel in the local case. Let $p: \mathcal{E} \to \mathcal{S}$ be a bounded $\mathcal{S}$-topos, where $\mathcal{S}$ has a natural number object. We recall that in B4.5.8 we constructed the 'measure topos' $M\mathcal{E}$, which has the property that geometric morphisms $\mathcal{F} \to M\mathcal{E}$ over $\mathcal{S}$ correspond to cocontinuous $\mathcal{S}$-indexed functors $\mathcal{E} \to \mathcal{S}$, for any bounded $\mathcal{S}$-topos $\mathcal{F}$. Suppose p is locally connected; then, for any $\mathcal{S}$-topos $q: \mathcal{F} \to \mathcal{S}$ and any cocontinuous $\mathcal{S}$-indexed functor $\theta: \mathcal{E} \to \mathcal{F}$, there is a unique $\mathcal{S}$-indexed natural transformation $\theta p^* \to q^*$ (since each object of $\mathcal{S}$ is an $\mathcal{S}$-indexed coproduct of copies of 1, and q^* preserves 1), and hence there is a unique $\mathcal{S}$-indexed natural transformation $\theta \to q^* p_!$. So $q^* p_!$ is terminal in $\mathfrak{LocPres}_{\mathcal{S}}(\mathcal{E}, \mathcal{F})$; equivalently, the section d of $M\mathcal{E} \to \mathcal{S}$ which corresponds to $p_!$ has the property that dq is terminal in $\mathfrak{Top}/\mathcal{S}(\mathcal{F}, M\mathcal{E})$. Thus we deduce that *if $\mathcal{E}$ is a locally connected bounded $\mathcal{S}$-topos, then $M\mathcal{E}$ is totally connected over $\mathcal{S}$*. In fact the converse is also true: for the canonical inclusion $\delta: \mathcal{E} \to M\mathcal{E}$ is $\mathcal{S}$-essential by the remark before B4.5.8, and hence so is its composite with $M\mathcal{E} \to \mathcal{S}$, even if we merely assume that the latter is locally connected. Note also that, if $\mathcal{E}$ is connected and locally connected over $\mathcal{S}$, then the topos $P\mathcal{E}$ of probability measures on $\mathcal{E}$ (cf. B4.5.12) is totally connected.

From 3.6.17(c), we may obtain a characterization of totally connected morphisms in terms of fibrations of sites: a morphism between bounded $\mathcal{S}$-toposes (where $\mathcal{S}$ has a natural number object) is totally connected iff it can be induced by a fibration of sites $(P, T): (\mathbb{D}, K) \to (\mathbb{C}, J)$ such that $\mathbb{C}$ and $\mathbb{D}$ are cartesian and P is a morphism of sites (i.e. it is cartesian and preserves covers) – cf. the remark after 3.1.21.

Lemma 3.6.18

(i) *A composite of totally connected morphisms is totally connected.*

(ii) *In a commutative triangle*

$$\begin{tikzcd} \mathcal{G} \arrow[r, "g"] \arrow[dr, "h"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

where h is totally connected and g is connected, f is totally connected. Moreover, if d is the right adjoint of h in $\mathfrak{Top}$, then gd is the right adjoint of f.

(iii) *Totally connected morphisms, and their right adjoints, are stable under pullback.*

(iv) *Given an inverse sequence of totally connected morphisms whose (***Cat***-enriched) limit exists in* $\mathfrak{T}\mathbf{op}$, *the legs of the limit cone are totally connected.*

(v) *Totally connected morphisms are orthogonal to grouplike morphisms (cf. 3.6.9).*

Proof (i) is immediate from either (i) or (iii) of 3.6.16.

(ii) If we define $f_! = h_! g^*$, then it is clearly cartesian; and it is left adjoint to f^* since g^* is full and faithful. So f has a right adjoint in $\mathfrak{T}\mathbf{op}$; but it is clearly connected, since g and h are.

(iii) and (iv) are proved in a similar manner to 3.6.7(iv) and (v); and (v) is exactly like 3.6.10. □

Similarly, for bounded toposes over a base, we have the dual of the notion of localization.

Theorem 3.6.19 *Given a bounded* $\mathcal{S}$*-topos* $p\colon \mathcal{E} \to \mathcal{S}$ *and a particular section* $s\colon \mathcal{S} \to \mathcal{E}$ *of* p*, there exists a bounded totally connected* $\mathcal{S}$*-topos* $p^s\colon \mathcal{E}^s \to \mathcal{S}$ *(called the* colocalization *of* $\mathcal{E}$ *at* s*) equipped with a geometric morphism* $\mathcal{E}^s \to \mathcal{E}$ *over* $\mathcal{S}$ *which carries the dense section of* $\mathcal{E}^s$ *into the chosen section* s*, and which induces an equivalence*

$$\mathfrak{B}\mathfrak{T}\mathbf{op}/\mathcal{S}\,(\mathcal{F},\mathcal{E}^s) \simeq (\mathfrak{B}\mathfrak{T}\mathbf{op}/\mathcal{S}\,(\mathcal{F},\mathcal{E}))/sq$$

for any bounded $\mathcal{S}$*-topos* $(q\colon \mathcal{F} \to \mathcal{S})$*.*

Proof $\mathcal{E}^s$ is constructed in the same way as the localization $\mathcal{E}_s$, except that we interchange the rôles of the domain and codomain maps $\mathbf{2} \pitchfork \mathcal{E} \rightrightarrows \mathcal{E}$. We have already observed in the proof of 3.6.12 that $d_0\colon \mathbf{2} \pitchfork \mathcal{E} \to \mathcal{E}$ is totally connected, its right adjoint being given by the diagonal Δ; so if we form the pullback

$$\begin{tikzcd}
\mathcal{E}^s \arrow[r, "\tilde{s}"] \arrow[d, "p^s"] & \mathbf{2} \pitchfork \mathcal{E} \arrow[d, "d_0"] \\
\mathcal{S} \arrow[r, "s"] & \mathcal{E}
\end{tikzcd}$$

and define $\mathcal{E}^s \to \mathcal{E}$ to be the composite $d_1 \tilde{s}$, then the stated properties of $\mathcal{E}^s$ all follow by arguments identical to those in the proof of 3.6.12. □

Suggestions for further reading: Bunge & Funk [197, 198], Johnstone & Moerdijk [545].

C4

LOCAL COMPACTNESS AND EXPONENTIABILITY

C4.1 Locally compact locales

Our principal goal in this chapter is to establish the characterization of exponentiable objects in the 2-category $\mathfrak{BTop}/\mathcal{S}$, for a general base topos $\mathcal{S}$, as those which are 'locally compact' in an appropriate sense. We have already made some observations about exponentiability in $\mathfrak{BTop}/\mathcal{S}$ in Section B4.3; and we shall assume that the reader of this chapter is familiar with them. In this and the next section we shall take our base topos to be **Set**, in order to simplify the notation; but virtually all that we do can be extended to a general base topos $\mathcal{S}$ with a natural number object, using an appropriate internal finiteness notion in $\mathcal{S}$ to interpret the finiteness conditions which appear below.

The notion of local compactness for locales was not considered in Section C1.5 (or in Chapter C3): this was a deliberate omission, so that we could begin the present chapter by discussing it, and thereby provide the motivation for the notion of continuous category which underlies the definition of local compactness for toposes.

First we need to recall the notions of continuous poset and continuous lattice.

Definition 4.1.1 Let A be a poset with directed joins (that is, a domain in the sense of 1.1.3).

(*a*) Given elements a and b of A, we say a is *way below* b (and write $a << b$) if, for every directed $S \subseteq A$ with $\bigvee S \geq b$, there exists $s \in S$ with $s \geq a$.

(*b*) We say A is a *continuous poset* if, for every $a \in A$, the set $\{b \in A \mid b << a\}$ is directed and has join a.

(*c*) We say A is a *continuous semilattice* (resp. *continuous lattice*) if it is a continuous poset and has finite meets (resp. joins).

Note that a poset with finite joins and directed joins has all joins, and is thus a complete lattice. It is also easy to see that the way-below relation is stable under finite joins, to the extent that they exist in A (that is, if $a_1 << b_1$ and $a_2 << b_2$, then $(a_1 \vee a_2) << (b_1 \vee b_2)$ provided the joins exist), so that in a complete lattice the directedness of the set $\{b \mid b << a\}$ is automatic. Another important property of the way-below relation on a continuous poset is the so-called *interpolation property*: if $a << b$, then there exists c satisfying $a << c$ and

$c << b$. To prove this, we simply note that the set $\{d \mid (\exists c)(d << c << b)\}$ is directed and has join b.

Lemma 4.1.2 *A continuous semilattice is a preframe. Hence a continuous lattice is a frame iff it satisfies the finite distributive law.*

Proof Let S be a directed subset of a continuous semilattice A, a an element of A. The inequality $\bigvee\{a \wedge s \mid s \in S\} \leq a \wedge \bigvee S$ is always true, so it suffices to show the reverse. But if $b << (a \wedge \bigvee S)$, then $b \leq a$ and $b \leq s$ for some $s \in S$, and hence $b \leq \bigvee\{a \wedge s \mid s \in S\}$; since $a \wedge \bigvee S$ is the join of all such b, the result follows.

The second assertion is immediate from the first, since a frame is precisely a preframe which is a (finitely) distributive lattice. □

We shall make extensive use of the Scott topology on continuous posets: recall that a subset U of a poset A with directed joins is said to be *Scott-open* if it is an upper set and inaccessible by directed joins (i.e. whenever the join of a directed set S is in U, then S meets U). The following facts about Scott-open subsets are easily established:

Lemma 4.1.3

(i) *For a poset A with directed joins, the set ΣA of all Scott-open subsets of A is a topology on A.*

(ii) *A subset U of a domain A is Scott-open iff its characteristic function $A \to \Omega$ is a morphism of domains.*

(iii) *If A is a continuous poset, then the sets $\Uparrow(a) = \{b \in A \mid a << b\}$ are Scott-open for all $a \in A$, and they form a base for the Scott topology on A.*

(iv) *Let $f: A \to B$ be a map between posets with directed joins. If f preserves directed joins, then it is continuous for the Scott topologies on A and B. The converse holds if B is continuous.*

Proof (i) and (ii) are trivial. For (iii), the openness of the sets $\Uparrow(a)$ follows from the interpolation property for $<<$, mentioned earlier; and they form a base because, if U is Scott-open and $a \in U$, we can find $b << a$ with $b \in U$, and hence $a \in \Uparrow(b) \subseteq U$. The first half of (iv) is trivial; for the second, suppose $S \subseteq A$ is directed and $b << f(\bigvee S)$ in B. Then $\bigvee S \in f^{-1}(\Uparrow(b))$, which is Scott-open if f is continuous, so there must exist $s \in S$ with $f(s) \geq b$. Hence

$$\bigvee\{f(s) \mid s \in S\} \geq \bigvee(\Downarrow(f(\bigvee S))) = f(\bigvee S);$$

the reverse inequality is trivial. □

There are two alternative characterizations of continuous posets, which will be of importance in what follows. For the first, we recall from 1.1.3 that an *ideal* in a poset is a subset which is (upwards) directed and also closed downwards; we write IA for the set of ideals of A, ordered by inclusion. The mapping which

sends an element a of A to the principal ideal $\downarrow(a)$ is an order-preserving map $A \to IA$.

Proposition 4.1.4 *The following conditions on a poset A are equivalent:*

(i) *A is continuous.*

(ii) *$\twoheaddownarrow : A \to IA$ has a left adjoint, which in turn has a left adjoint.*

(iii) *A is a retract in* **Dom** *of a poset of the form IB.*

(iv) *A is a retract in* **Dom** *of a continuous poset.*

Proof (i) $\Rightarrow$ (ii): If A has directed joins, then the mapping $\bigvee : IA \to A$ which sends an ideal to its join is easily seen to be left adjoint to $\downarrow$. The definition of the way-below relation says that the left adjoint of $\bigvee$, if it exists, necessarily sends an element a to $\twoheaddownarrow(a) = \{b \in A \mid b << a\}$; the definition of continuity says that the map $\twoheaddownarrow$ does indeed take values in IA, and that the unit of the adjunction exists (the existence of the counit is automatic from the definition of $<<$).

(ii) $\Rightarrow$ (iii): Reversing the first part of the previous argument, if $\downarrow$ has a left adjoint $\bigvee$, then the latter necessarily sends an ideal to its join in A; hence A has directed joins, since the join of any directed set coincides with that of the ideal which is its downward closure. Now $\bigvee$ and its left adjoint both preserve any joins which exist; and they express A as a retract of IA, since $\downarrow$ is injective and hence $\bigvee$ is surjective.

(iii) $\Rightarrow$ (iv): We have to show that a poset of the form IB is continuous. But IB certainly has directed joins, since they are just unions; and any principal ideal $\downarrow(b)$ is inaccessible by directed joins, i.e. satisfies $\downarrow(b) << \downarrow(b)$. Since any ideal I is the (directed) union of the principal ideals which it contains, it follows easily that $J << I$ iff J is contained in $\downarrow(b)$ for some $b \in I$. It is now immediate that $\twoheaddownarrow(I)$ is directed and has I as its union.

(iv) $\Rightarrow$ (i): Suppose B is continuous, and let $i : A \to B$, $r : B \to A$ be directed-join-preserving maps with $ri = 1_A$. (It is then clear that A has directed joins.) Given $a \in A$, we claim first that $r(b) << a$ in A whenever $b << i(a)$ in B: for if $S \subseteq A$ is directed with $\bigvee S \geq a$, then $b \leq i(s)$ for some $s \in S$, whence $r(b) \leq ri(s) = s$. Since $a = ri(a)$ is the join of all $r(b)$ with $b << i(a)$, it follows as in the previous paragraph that $a' << a$ in A iff $a' \leq r(b)$ for some $b << i(a)$, and hence that A is continuous. $\square$

Remark 4.1.5 We note that, if A is expressible as a retract in **Dom** of a poset of the form IB, then it is expressible as a coadjoint retract (in the sense of B1.1.9(c)) of such a poset; i.e. we can find $i : A \to IC$ and $r : IC \to A$ such that i is left adjoint to r and the unit $1_A \to ri$ is an equality. Equivalently, if we start from the full subcategory of **Dom** whose objects are those of the form IB, then splitting all the idempotents in this category has no more effect than splitting the idempotent *comonads*, i.e. those idempotents $e : IB \to IB$ satisfying $e \leq 1_{IB}$. Results of this type will form a recurring theme throughout this chapter.

If the remarks in the previous paragraph seem reminiscent of B1.1.15, that is no accident: of course, what is going on here is that we are dealing with a KZ-monad on the (locally ordered) 2-category $\mathfrak{Poset}$, whose functor part is given by $A \mapsto IA$, whose unit is the principal ideals map $\downarrow : A \to IA$, and whose multiplication is the join map $I(IA) \to IA$. The algebras for this monad are exactly the domains; so the equivalence of (ii) and (iii) in 4.1.4 is more or less a direct transcription of B1.1.15 in this particular case.

For the second characterization (which was Scott's original justification for studying continuous lattices [1102]), we restrict our attention to lattices. Note that, in the characterization of 4.1.4(iii), continuous lattices are just the retracts in **Dom** of posets of the form IB where B is a join-semilattice.

Proposition 4.1.6 *For a complete lattice A, the following are equivalent:*

(i) *A is continuous.*

(ii) *$(A, \Sigma A)$ is a sober space, and is injective with respect to subspace inclusions in the category* **Sob**.

(iii) *A is the set of points of a locale X which is injective with respect to inclusions in* **Loc**, *ordered by the 'specialization ordering' (that is, the canonical partial ordering on maps $1 \to X$ in* **Loc**, *as defined in* 1.1.6*).*

(iv) *$(A, \Sigma A)$ is a retract of a power of the Sierpiński space S (cf.* 1.1.4*).*

Proof (i) $\Rightarrow$ (ii) and (iii): First consider the case when $A = IB$ for a join-semilattice B. Then a Scott-open subset $U \subseteq A$ is determined by the set $\{b \in B \mid \downarrow(b) \in U\}$, since any element of A is a directed join of principal ideals; but the latter set may be any upper subset of B, so ΣA is isomorphic to the lattice UB of all upper sets in B. But the latter may also be regarded as the lattice $L(B^{\mathrm{op}})$ of lower sets in the meet-semilattice B^{op}, and we saw in 1.1.3 that this is the free frame generated by B^{op}. Thus a frame homomorphism $\phi : UB \to \Omega$ corresponds to a meet-semilattice homomorphism $B^{\mathrm{op}} \to \Omega$; but the latter is the same thing as the characteristic function of a filter of B^{op}, i.e. of an ideal of B. So we may identify $(A, \Sigma A)$ with the space of points of the locale X defined by $\mathcal{O}(X) = UB$. In particular, it is sober, and the locale X is spatial.

To verify that it is injective, we observe that for any locale Y (in particular for any sober space), a continuous map $f : Y \to X$ corresponds to a frame homomorphism $f^* : L(B^{\mathrm{op}}) \cong \mathcal{O}(X) \to \mathcal{O}(Y)$ and hence to a meet-semilattice homomorphism $B^{\mathrm{op}} \to \mathcal{O}(Y)$. Now if Y is a sublocale of Z, then $\mathcal{O}(Y)$ is a retract of $\mathcal{O}(Z)$ in **mSLat**, since frame homomorphisms and their right adjoints are both meet-semilattice homomorphisms; hence we may extend any continuous map $Y \to X$ to a map $Z \to X$. Thus we have verified that (ii) and (iii) hold whenever A is of the form IB.

The general case follows easily from 4.1.4(iii): for if A is a retract of IB by maps preserving directed joins, the latter are continuous with respect to the Scott topologies by 4.1.3(iv). But any retract of a sober space is sober, since

Sob is reflective in **Sp** and hence closed under equalizers; and a retract of an injective object is injective in any category.

(ii) ⇒ (iv): Any sober space (indeed, any T_0-space) X may be embedded as a subspace of a power S^B of Sierpiński space: just take B to be the set $\mathcal{O}(X)$, and map $x \in X$ to the point whose Uth coordinate is 1 iff $x \in U$. If X is injective, this embedding must have a retraction.

(iii) ⇒ (iv): Similarly, we saw in 1.1.4 that the free objects of **Frm** are exactly the topologies of powers of Sierpiński space. It follows that the projectives (with respect to regular epimorphisms) in **Frm** are exactly the retracts of such powers (in particular, they are all spatial, which explains the equivalence of (ii) and (iii)).

(iv) ⇒ (i): First we observe that the product topology on any power S^B of Sierpiński space coincides with the Scott topology for the pointwise ordering on S^B (where S is given the obvious ordering $0 < 1$). But $S^B \cong PB \cong I(KB)$ is clearly a continuous lattice. If $(A, \Sigma A)$ is a retract of S^B, then the inclusion and retraction must preserve directed joins, since they are continuous for the Scott topologies; so A is continuous by 4.1.4(iv). □

We remark that it follows from 4.1.6 that, for any injective sober space X, the specialization ordering on points of X is a complete-lattice ordering. (Recall that it always has directed joins, by 1.1.6; so the force of this observation is that it has finite joins.)

In the lattice $\mathcal{O}(X)$ of open subsets of a topological space X, it is easy to see that the relation $U \ll V$ holds if there is a compact set $K \subseteq X$ with $U \subseteq K \subseteq V$. (In particular, $U \ll U$ holds iff U itself is compact.) Hence if X is a locally compact space (i.e. one in which every open set is covered by the interiors of its compact subsets), then $\mathcal{O}(X)$ is a continuous lattice. Conversely, it can be shown using the axiom of choice that if a continuous lattice A is distributive, then the locale X defined by $\mathcal{O}(X) = A$ is spatial, and its space of points is locally compact. (We shall not give the proof here; see [452] for the details.) We therefore define a *locally compact locale* to be one whose frame of opens is a continuous lattice.

Given a nucleus j on a frame $\mathcal{O}(X)$, it is easy to see that the sublocale corresponding to the fixset $\mathcal{O}(X)_j$ is compact iff the filter $F = \{U \in \mathcal{O}(X) \mid j(U) = X\}$ is Scott-open. In general, not every filter in a frame is representable as the filter of elements mapped to the top element by a nucleus; but it turns out that every Scott-open filter is so representable (see [526]). Using this, plus the result of [45] that distributive continuous lattices have 'enough Scott-open filters', it is possible to give an alternative characterization of locally compact locales as those which have 'enough compact sublocales', which looks more like the usual definition of local compactness for spaces. However, we shall not need this characterization here; and it is in any case non-constructive (the second ingredient mentioned above requires the axiom of countable dependent choices), so we cannot use it when dealing with internal locales in a topos.

For future reference, we note a couple of stability properties of local compactness.

Lemma 4.1.7

(i) *A locally closed sublocale of a locally compact locale is locally compact.*

(ii) *If $f: X \to Y$ is a triquotient map in the sense of 3.2.7 (for example, an open or a proper surjection) and X is locally compact, then so is Y.*

Proof (i) Since a locally closed sublocale is an open sublocale of its closure, it suffices to prove separately that local compactness is inherited by open and by closed sublocales. But each of these is easy, since principal ideals and principal filters in a continuous lattice are continuous (either directly from the definition, or using 4.1.4(iv)).

(ii) We can express $\mathcal{O}(Y)$ as a retract of $\mathcal{O}(X)$ by maps preserving directed joins. So this is immediate from 4.1.4(iv). □

Local compactness is also inherited by finite products in **Loc**: it is not hard to prove this directly from the definition, but we do not need to do so, since it follows from 4.1.9 below and A1.5.1(i). On the other hand, we shall give a constructive proof that spatial and localic products coincide if one of the factors is locally compact. (For more information on when spatial and localic products coincide, see [478] and [978].)

Lemma 4.1.8 *Suppose X and Y are spatial locales, and X is locally compact. Then the locale product $X \times Y$ is spatial.*

Proof Let W and W' be opens in $X \times Y$ such that every point of W belongs to W'; we must show that $W \leq W'$, for which it is enough to show that every open rectangle $U \times V \leq W$ satisfies $U \times V \leq W'$. For any open $V' \leq Y$, let $U_{V'} = \bigcup\{U' \in \mathcal{O}(X) \mid U' \times V' \leq W'\}$; and for any y in V, let $U_y = \bigcup\{U_{V'} \mid V' \in \mathcal{O}(Y), y \in V'\}$. For each $(x, y) \in U \times V$, we have $(x, y) \in W'$, so there must be a rectangular neighbourhood $U' \times V'$ of (x, y) contained in W', and hence $x \in U_y$. Since X is spatial, U is covered by its points, and so this implies $U \leq U_y$ for each $y \in V$; but the join in the definition of U_y is directed, and hence for each $U' << U$ we may conclude that every $y \in V$ has a neighbourhood V' such that $U' \times V' \leq W'$. Since V is also covered by its points, this in turn implies that we have $U' \times V \leq W'$ for each $U' << U$; finally, local compactness of X implies that we have $U \times V \leq W'$. □

A well-known application of local compactness in 'classical' topology is that locally compact spaces are exactly those for which the compact-open topology on function spaces 'works well' (cf. [367]): in the terminology introduced in Section A1.5, they are exactly the exponentiable objects of **Sp**. The corresponding result for locales was first proved by M. Hyland [461]:

Theorem 4.1.9 *A locale is exponentiable in* **Loc** *iff it is locally compact.*

Proof First suppose X is exponentiable. Since the functor $(-) \times X : \mathbf{Loc} \to \mathbf{Loc}$ preserves inclusions (as they are just regular monomorphisms), its right adjoint $(-)^X$ preserves injectives; in particular, the exponential S^X must be injective, where S is the Sierpiński space. But points of S^X correspond bijectively to locale maps $X \to S$, and hence (since $\mathcal{O}(S)$ is the free frame on one generator) to elements of $\mathcal{O}(X)$; we claim that this correspondence maps the specialization ordering on points of S^X to the given ordering on the frame $\mathcal{O}(X)$, so that the latter must be a continuous lattice by 4.1.6.

To verify the claim, note that the correspondence sends a point $p: 1 \to S^X$ to $q^*(U) \in \mathcal{O}(X)$, where q is the composite

$$X \cong 1 \times X \xrightarrow{p \times 1} S^X \times X \xrightarrow{\mathrm{ev}} S$$

and U is the generator of $\mathcal{O}(S)$ when the latter is regarded as the free frame on one generator. Clearly $p_1 \leq p_2$ implies $q_1 \leq q_2$ and hence $q_1^*(U) \leq q_2^*(U)$. Conversely, if $V_1 \leq V_2$ in $\mathcal{O}(X)$, consider the element

$$W = (S \otimes V_1) \cup (U \otimes V_2)$$

of $\mathcal{O}(S \times X) \cong \mathcal{O}(S) \otimes \mathcal{O}(X)$; this element corresponds to a locale map $S \times X \to S$, whose transpose $w: S \to S^X$ has the property that $w \cdot 0$ and $w \cdot 1$ are the points corresponding to V_1 and V_2 respectively. (Here 0 and 1 are the two points of S.) But $0 \leq 1$ in the specialization order on S, so $w \cdot 0 \leq w \cdot 1$.

Conversely, suppose X is locally compact. Let Y be the locale defined by $\mathcal{O}(Y) = \Sigma(\mathcal{O}(X))$; then Y is an injective locale by 4.1.6(iii), and it is clearly the only possible candidate for the exponential S^X in $\mathbf{Loc}$. We first verify that it does indeed have the universal property of an exponential S^X. The evaluation map $\mathrm{ev}: Y \times X \to S$ must correspond to an element of $\mathcal{O}(Y \times X) \cong \Sigma(\mathcal{O}(X)) \otimes \mathcal{O}(X)$; for this we take

$$E = \bigcup \{(\uparrow(U) \otimes U) \mid U \in \mathcal{O}(X)\} .$$

Given a locale Z and a map $h: Z \times X \to S$, corresponding to an open $H \in \mathcal{O}(Z \times X) \cong \mathcal{O}(Z) \otimes \mathcal{O}(X)$, we define $\overline{h}: Z \to Y$ by

$$\overline{h}^*(V) = \bigcup \{W \in \mathcal{O}(Z) \mid (\exists U \in V)(W \otimes U \leq H)\} .$$

(It is straightforward to verify that $\overline{h}^*$ is a frame homomorphism.) We claim that the diagram

$$\begin{tikzcd} Z \times X \arrow[r, "\overline{h} \times 1"] \arrow[dr, "h"'] & Y \times X \arrow[d, "\mathrm{ev}"] \\ & S \end{tikzcd}$$

commutes; for this, it clearly suffices to show that $(\overline{h} \times 1)^*(E) = H$. But by definition, we have

$$(\overline{h} \times 1)^*(E) = \bigcup\{\overline{h}^*(\Uparrow(U)) \otimes U \mid U \in \mathcal{O}(X)\}$$

and

$$\overline{h}^*(\Uparrow(U)) \otimes U = \bigcup\{W \otimes U \mid (\exists U' \in \Uparrow(U))(W \otimes U' \leq H)\}\,.$$

The union of the right-hand side of the latter equation is clearly below H, since $U << U'$ implies $W \otimes U \leq W \otimes U'$; so $(\overline{h} \times 1)^*(E) \leq H$. But H can be written as a union of 'open rectangles' $W \otimes U$, since these generate $\mathcal{O}(Z \times X)$ as a complete join-semilattice, and hence by continuity of $\mathcal{O}(X)$ it can also be written as a join of rectangles $W \otimes U$ for which there exists U' with $U << U'$ and $W \otimes U' \leq H$. So the reverse inequality also holds.

We must next show that if $k\colon Z \to Y$ is any locale map which satisfies $\mathrm{ev}(k \times 1) = h$, then $k = \overline{h}$. Since the Scott-open sets of the form $\Uparrow(U)$ form a base for $\Sigma(\mathcal{O}(X))$ by 4.1.4(iii), it suffices to show that k^* and $\overline{h}^*$ agree on all such sets. But we have

$$H = (k \times 1)^*E = \bigcup\{k^*(\Uparrow(U)) \times U \mid U \in \mathcal{O}(X)\},$$

and hence for any U' with $U << U'$ we have

$$\overline{h}^*(\Uparrow(U)) \geq k^*(\Uparrow(U'))$$

by the definition of $\overline{h}^*$. But $\Uparrow(U)$ is the union of all $\Uparrow(U')$ with $U << U'$, by the interpolation property; so it follows that we have $k^*(\Uparrow(U)) \leq \overline{h}^*(\Uparrow(U))$ for all U.

For the reverse inequality, suppose $U << U'$ and $W \in \mathcal{O}(Z)$ are such that $W \times U' \leq H$. Then $W \times V'$ is covered by the opens $k^*(\Uparrow(V)) \times V$, $V \in \mathcal{O}(X)$. Now the set $\{V \in \mathcal{O}(X) \mid k^*(\Uparrow(V)) \geq W\}$ is directed, since we have $k^*(\Uparrow(V_1 \cup V_2)) = k^*(\Uparrow(V_1) \cap \Uparrow(V_2)) = k^*(\Uparrow(V_1)) \cap k^*(\Uparrow(V_2))$, and it covers U', so there exists $V \geq U$ with $k^*(\Uparrow(V)) \geq W$. But this implies that $\overline{h}^*(\Uparrow(U')) \leq k^*(\Uparrow(U))$, whence $\overline{h}^*(\Uparrow(U)) \leq k^*(\Uparrow(U))$ follows as before.

Having thus constructed the exponential S^X, we may clearly now define $(S^A)^X$, where A is any set, as $(S^X)^A$ (that is, the product in **Loc** of A copies of S^X). And any continuous map $f\colon S^A \to S^B$ gives rise to a map $f^X : (S^A)^X \to (S^B)^X$, namely the transpose of

$$(S^A)^X \times X \xrightarrow{\ \mathrm{ev}\ } S^A \xrightarrow{\ \ f\ \ } S^B;$$

that is, $(-)^X$ becomes a functor on the full subcategory of **Loc** whose objects are the powers of S. But any locale Y is a sublocale of a power of S, as we observed in the proof of 4.1.6; and locale inclusions are regular monomorphisms, so we can express Y as the equalizer of a pair of maps $S^A \rightrightarrows S^B$. Since the functor

$(-)^X$ is a right adjoint, it should preserve equalizers; so we have merely to define Y^X to be the equalizer of the corresponding pair of maps $(S^A)^X \rightrightarrows (S^B)^X$, and to verify that this has the right universal property. $\square$

We remark that it is possible to give a direct construction of the exponential Y^X for an arbitrary Y and locally compact X, by constructing a presentation for its frame of opens from the elements of $\mathcal{O}(Y)$ and $\mathcal{O}(X)$: this is done in Hyland's original paper [461]. However, the fact that the existence of the particular exponential S^X is sufficient to construct Y^X for an arbitrary Y is of interest in its own right – note that we have seen the topos-theoretic version of this locale-theoretic fact in B4.3.1.

To conclude this section, we briefly discuss stable local compactness. We have seen that the way-below relation in a complete lattice is always stable under finite joins; but it need not be stable under finite meets, even if the lattice is continuous and distributive:

Example 4.1.10 As in 3.4.1(a), let X be the space obtained from two copies of the closed unit interval $[0, 1]$ by identifying the two copies of the open interval $(0, 1)$. Then the two copies of $[0, 1]$ are open in X and compact, so they define elements U and V of $\mathcal{O}(X)$ satisfying $U \ll U$ and $V \ll V$. But $U \cap V \cong (0, 1)$ is not compact, so we do not have $(U \cap V) \ll (U \cap V)$.

We say a continuous (semi)lattice A is *stably continuous* if the way-below relation on A is stable under finite meets (including the empty meet; i.e. the top element 1_A is compact); and we call a locale X *stably locally compact* if $\mathcal{O}(X)$ is a stably continuous lattice. (If we merely wish to assert that $\ll$ is stable under binary meets, without assuming compactness of the top element, we shall say that X is *locally stably compact.*)

Lemma 4.1.11 *Any compact Hausdorff locale is stably locally compact.*

Proof We recall from 3.2.10(ii) that a compact Hausdorff locale X is regular; that is, every open U is the join of the opens V such that there exists an open W with $V \cap W = 0$ and $U \cup W = X$ – as in 1.2.17, we denote this relation between U and V by $V \lhd U$. Since the relation $\lhd$ is clearly stable under finite meets, it suffices to show that it coincides with $\ll$.

But $\lhd$ is also stable under finite joins, so the join appearing in the definition of regularity is directed; thus in a regular locale the relation $V \ll U$ implies $V \leq W$ for some $W \lhd U$, and hence $V \lhd U$. And any closed sublocale of a compact locale is compact, so in a compact locale the relation $V \lhd U$ implies $V \ll U$. So the result is established. $\square$

For future reference, we also note:

Lemma 4.1.12 *Let X be a compact Hausdorff locale. Then a sublocale of X is locally compact iff it is locally closed.*

Proof One direction is immediate from 4.1.11 and 4.1.7(i). Conversely, suppose Y is a locally compact sublocale of X. By replacing X by the closure of Y, we may assume Y is dense in X; we then have to prove that it is open. We first note that if $V << Y$ in $\mathcal{O}(Y)$, then the closure $\overline{V} = \complement(\neg V)$ of V in Y is compact; for if we have a directed family of opens U_i with $\bigcup_i U_i \geq \overline{V}$, then $\bigcup_i (U_i \cup \neg V) = Y$, whence if $V << W << Y$ we have $U_i \cup \neg V \geq W$ for some i. But Y is regular since it is a sublocale of X; so, as in the proof of 4.1.11, the relation $V << W$ implies $V \triangleleft W$, and hence $U_i \cup \neg V \geq W$ implies $U_i \geq \overline{V}$.

Now by 3.2.10(i) we know that all compact sublocales of X are closed, so $\overline{V}$ must actually be closed in X. Suppose it equals $\complement U$ where $U \in \mathcal{O}(X)$, and let $\tilde{V} \in \mathcal{O}(X)$ be such that $\tilde{V} \cap Y = V$. Then $\tilde{V} \cap Y \cap U = \emptyset$, whence $\tilde{V} \cap U = \emptyset$ since Y is dense, so $\tilde{V} \leq \complement U = \overline{V}$. But $\overline{V} \leq Y$ by definition, so $\tilde{V} \leq Y$, and thus $V = \tilde{V}$ is open in X. Since Y is locally compact, we can express it as a union of opens V which satisfy $V << Y$ and are therefore open in X; so Y is open in X. □

We recall from 1.1.3 that if B is a distributive lattice, then its ideal lattice IB is (distributive, and hence) a frame; as in 2.4.3, we call a locale X *coherent* if $\mathcal{O}(X) \cong IB$ for some distributive lattice B. (Of course, B is determined up to isomorphism by X: it is the lattice of compact open sublocales of X.)

Proposition 4.1.13 *For a locale X, the following are equivalent:*

(i) *X is stably locally compact.*

(ii) *X is a retract in* **Loc** *of a coherent locale.*

(iii) *X is a coadjoint retract in* **Loc** *of a coherent locale.*

Proof (i) ⇒ (iii): Suppose X is stably locally compact. Then, as in 4.1.4, we know that $\mathcal{O}(X)$ is a retract of $I(\mathcal{O}(X))$ by maps $\twoheaddownarrow : \mathcal{O}(X) \to I(\mathcal{O}(X))$ and $\bigvee : I(\mathcal{O}(X)) \to \mathcal{O}(X)$ preserving arbitrary joins; we claim that these maps also preserve finite meets, i.e. they are frame homomorphisms. For $\bigvee$, this is obvious, since it is right adjoint to $\twoheaddownarrow$; and the preservation of finite meets by $\twoheaddownarrow$ is exactly the definition of stability for $<<$. Thus we have expressed X as a retract of the coherent locale Y defined by $\mathcal{O}(Y) = I(\mathcal{O}(X))$; moreover, since $\twoheaddownarrow$ is left adjoint to $\bigvee$, it is actually a coadjoint retract as defined in B1.1.9(c).

(iii) ⇒ (ii) is trivial. For (ii) ⇒ (i), we show that coherent locales are stably locally compact and that retracts of stably locally compact locales are stably locally compact. The first of these is easy, since in the frame IB (for a distributive lattice B) we have $I << J$ iff there exists a principal ideal $\downarrow(b)$ with $I \subseteq \downarrow(b) \subseteq J$; and the set of principal ideals is closed under finite meets. For the second, suppose X is a retract of Y by locale maps $i: X \to Y$ and $r: Y \to X$; then, from the proof of (iv) ⇒ (i) in 4.1.4, we know that $U << V$ in $\mathcal{O}(X)$ iff $U \leq i^*(W)$ for some $W << r^*(V)$ in $\mathcal{O}(Y)$. Since Y is stably locally compact and both r^* and i^* preserve finite meets, it is easy to see that this relation is stable under finite meets. □

We remark that, by essentially the same proof, locally stably compact locales are exactly the retracts in **Loc** of *locally coherent locales*, the latter being defined as those whose frames are of the form IB where B is a 'distributive pre-lattice'; i.e. it satisfies all the axioms for a distributive lattice, except that it may lack a top element. (Given a locally stably compact locale X, we may express $\mathcal{O}(X)$ as a retract of IB where $B = \{U \in \mathcal{O}(X) \mid U << X\}$.)

We may now prove a result promised in Section C3.4:

Corollary 4.1.14 *Stably locally compact locales (in particular, compact Hausdorff locales) are strongly compact.*

Proof If X is a coherent locale, say $\mathcal{O}(X) \cong IB$ for a distributive lattice B, then the Comparison Lemma 2.2.3 yields an equivalence $\mathbf{Sh}(X) \simeq \mathbf{Sh}(B, C)$, where C is the coherent coverage on B considered as a coherent category (cf. A2.1.11(b)). So X is strongly compact by 3.4.1(c). But if X is a coadjoint retract of a strongly compact locale Y, consider the diagram

$$\begin{tikzcd} \mathbf{Sh}(X) \arrow[rr, shift left, "i_*"] \arrow[dr, "\gamma_*"'] & & \mathbf{Sh}(Y) \arrow[ll, shift left, "r_*"] \arrow[dl, "\delta_*"] \\ & \mathbf{Set} & \end{tikzcd}$$

where i and r are the inclusion and the retraction. Given a filtered diagram in $\mathbf{Sh}(X)$ with vertices $(A_j \mid j \in J)$, we have

$$\begin{aligned} \varinjlim_J \gamma_*(A_j) &\cong \varinjlim_J \delta_* i_*(A_j) \cong \delta_*(\varinjlim_J i_*(A_j)) \cong \gamma_* r_*(\varinjlim_J i_*(A_j)) \\ &\cong \gamma_*(\varinjlim_J r_* i_*(A_j)) \cong \gamma_*(\varinjlim_J A_j) \ , \end{aligned}$$

where the fourth isomorphism uses the fact that r_* is left adjoint to i_*. So X is strongly compact. □

Another application of 4.1.13, sometimes known as 'Joyal's Lemma' (see [520, III 1.11]), is of use in various contexts.

Corollary 4.1.15

(i) *A locale X is stably locally compact iff it is injective with respect to flat inclusions of sublocales.*

(ii) *Compact Hausdorff locales are orthogonal to flat inclusions; i.e. if X is compact Hausdorff and $i: Y \rightarrowtail Z$ is a flat inclusion of locales, then any locale morphism $Y \to X$ extends uniquely to a morphism $Z \to X$.*

Proof (i) First suppose X is stably locally compact. Recall from 1.1.16(d) that a locale inclusion $i: Y \rightarrowtail Z$ is said to be flat if $i_* : \mathcal{O}(Y) \to \mathcal{O}(Z)$ preserves finite joins – equivalently, if it is a lattice homomorphism. But the assignment

$A \mapsto I(A)$ is a functor **DLat** $\to$ **Frm** – indeed, it is left adjoint to the forgetful functor, as we saw in 1.1.3. So, given a frame homomorphism $f^* \colon \mathcal{O}(X) \to \mathcal{O}(Y)$, we may define $g^* \colon \mathcal{O}(X) \to \mathcal{O}(Z)$ to be the composite frame homomorphism

$$\mathcal{O}(X) \xrightarrow{\downarrow} I\mathcal{O}(X) \xrightarrow{I(f^*)} I\mathcal{O}(Y) \xrightarrow{I(i_*)} I\mathcal{O}(Z) \xrightarrow{\bigvee} \mathcal{O}(Z) .$$

We must verify that $i^* g^* = f^*$; but

$$\begin{aligned} i^* g^* &= i^* \bigvee\nolimits_Z I(i_*) I(f^*)\downarrow \\ &= \bigvee\nolimits_Y I(i^*) I(i_*) I(f^*)\downarrow \\ &= \bigvee\nolimits_Y I(f^*)\downarrow \\ &= f^* \bigvee\nolimits_X \downarrow \\ &= f^*, \end{aligned}$$

where the last step uses 4.1.13, and the others are instances of the functoriality of I or the naturality of $\bigvee$.

For the converse, we recall from 1.1.16(*e*) that any locale X is a flat sublocale of a coherent locale; so if X is injective with respect to flat inclusions it must be a retract of a coherent locale.

(ii) Since compact Hausdorff locales are stably locally compact by 4.1.11, the existence of the extension follows from (i). For the uniqueness, we use the fact that if X is Hausdorff then the equalizer of any two locale morphisms $Z \rightrightarrows X$ is a closed sublocale of Z, because it is a pullback of the diagonal $X \rightarrowtail X \times X$; but flat inclusions are dense, so if the equalizer contains Y then it must be the whole of Z. □

Suggestions for further reading: Banaschewski [45], Gierz et al. [404], Hofmann [450], Hofmann & Lawson [452], Hyland [461], Johnstone [520, 523], Niefield [894], Scott [1102].

C4.2 Continuous categories

In this section, we develop an analogue for categories of the notion of continuous poset which we recalled in the previous section. For this purpose, it is most convenient to start from the characterization provided by 4.1.4: the fact that continuous posets are the retracts of posets of the form IB by maps preserving directed joins. Obviously, when we generalize from posets to categories, we need to replace directed joins by filtered colimits; and the analogue of the ideal completion IB (the poset obtained by freely adding directed joins to a poset B) is

the *inductive completion* Ind-$\mathcal{C}$ of a category $\mathcal{C}$. We therefore begin by reviewing this notion.

Throughout this section, we shall work explicitly with **Set**-based categories; but all our results may be reinterpreted in the context of $\mathcal{S}$-indexed categories for an arbitrary base topos $\mathcal{S}$, using the techniques developed in Part B, and when we come to apply the results of this section to toposes in the subsequent sections of this chapter, we shall in fact do so at that level of generality.

If $\mathcal{C}$ is locally small, then it is fully embedded (by the Yoneda embedding) in the functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, and the latter may be viewed as the free cocompletion of $\mathcal{C}$, as we saw in B2.5.8. If we merely wish to adjoin filtered colimits, we may restrict our attention to the full subcategory $\widehat{\mathcal{C}}$ of those functors which are expressible as (small) filtered colimits of representable functors. (If $\mathcal{C}$ is small, the objects of $\widehat{\mathcal{C}}$ are just the $\mathcal{C}$-torsors as defined in B3.2.3, i.e. those functors $F\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ for which the domain $\mathcal{F}$ of the corresponding discrete fibration $\mathcal{F} \to \mathcal{C}$ is filtered.) We note that this category has the following properties:

(i) There is a full and faithful functor $y\colon \mathcal{C} \to \widehat{\mathcal{C}}$.

(ii) $\widehat{\mathcal{C}}$ has filtered colimits, and every object of $\widehat{\mathcal{C}}$ is expressible as a (small) filtered colimit of objects in the image of y.

(iii) The objects in the image of y are *finitely presentable* as defined in D2.3.1(a); that is, the functor $\widehat{\mathcal{C}}\,(y(A), -)$ preserves filtered colimits for every $A \in \mathrm{ob}\, \mathcal{C}$.

(These three properties should be compared with those which we used to characterize the regularization $\mathbf{Reg}(\mathcal{C})$ of a cartesian category $\mathcal{C}$ in Section A1.3.) The first two properties are obvious from the definition of $\widehat{\mathcal{C}}$; for the third, we note that as an object of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ the functor represented by $y(A)$ preserves all small colimits, by the Yoneda lemma, and the inclusion $\widehat{\mathcal{C}} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ creates filtered colimits.

In fact properties (i)–(iii) suffice to characterize $\widehat{\mathcal{C}}$ up to (weak) equivalence, and provide the key to an alternative definition which is in many ways more convenient to use. To see this, note that condition (ii) says that we might as well take the objects of $\widehat{\mathcal{C}}$ to be small filtered diagrams in $\mathcal{C}$, i.e. functors $D\colon I \to \mathcal{C}$ where I is a small filtered category. (We shall normally denote such a diagram by $(A_i \mid i \in I)$, where A_i is the object $D(i)$ of $\mathcal{C}$, leaving the reader's imagination to supply names for the edges of the diagram.) For two such objects $(A_i \mid i \in I)$ and $(B_j \mid j \in J)$, the class of morphisms between them must satisfy

$$\begin{aligned}\widehat{\mathcal{C}}\,((A_i \mid i \in I), (B_j \mid j \in J)) &\cong \varprojlim_{i\in I}\widehat{\mathcal{C}}\,(y(A_i), (B_j \mid j \in J)) \\ &\cong \varprojlim_{i\in I}\varinjlim_{j\in J}\widehat{\mathcal{C}}\,(y(A_i), y(B_j)) \\ &\cong \varprojlim_{i\in I}\varinjlim_{j\in J}\mathcal{C}\,(A_i, B_j)\end{aligned}$$

using conditions (iii) and (i), plus the fact that $(A_i \mid i \in I)$ is intended to be the colimit in $\widehat{\mathcal{C}}$ of the $y(A_i)$. Accordingly, we now define Ind-$\mathcal{C}$ to be the category whose objects are all small filtered diagrams in $\mathcal{C}$, and whose morphisms are defined by the 'double-limit' formula in the last line of the above display: in other words, a morphism $(A_i \mid i \in I) \to (B_j \mid j \in J)$ is a family $(\phi_i \mid i \in I)$, where each ϕ_i is an equivalence class of morphisms $f \colon A_i \to B_j$ (two morphisms $f \colon A_i \to B_j$ and $f' \colon A_i \to B_{j'}$ being equivalent iff there exist morphisms $j \to j''$ and $j' \to j''$ in J such that

$$\begin{tikzcd} A_i \arrow[r, "f"] \arrow[d, "f'"] & B_j \arrow[d] \\ B_{j'} \arrow[r] & B_{j''} \end{tikzcd}$$

commutes), subject to the obvious compatibility conditions. The composite of two morphisms

$$(A_i \mid i \in I) \xrightarrow{(\phi_i \mid i \in I)} (B_j \mid j \in J) \xrightarrow{(\psi_j \mid j \in J)} (C_k \mid k \in K)$$

is $(\chi_i \mid i \in I)$, where χ_i is the element of $\varinjlim_{k \in K} \mathcal{C}(A_i, C_k)$ containing all the composites gf such that $f \colon A_i \to B_j$ is in ϕ_i and $g \colon B_j \to C_k$ is in ψ_j. The embedding $y \colon \mathcal{C} \to$ Ind-$\mathcal{C}$ sends an object A to itself regarded as a diagram over the terminal category $\mathbf{1}$; it is clear that this is indeed full and faithful.

Remark 4.2.1 If $\mathcal{C}$ is a preorder, then so is Ind-$\mathcal{C}$: for if the hom-sets $\mathcal{C}(A_i, B_j)$ each have at most one element, then so do the filtered colimits $\varinjlim_{j \in J} \mathcal{C}(A_i, B_j)$, and hence so does the double limit $\varprojlim_{i \in I} \varinjlim_{j \in J} \mathcal{C}(A_i, B_j)$. In particular, if B is a (small) poset, then Ind-B is equivalent to the ideal completion IB: in one direction, the equivalence maps a filtered diagram $(a_i \mid i \in I)$ to the ideal $\{b \in B \mid (\exists i \in I)(b \leq a_i)\}$, and in the other it sends an ideal $J \subseteq B$ to the inclusion $J \to B$ regarded as a filtered diagram in B. This ensures that the definition of continuity for categories, which we are about to give, genuinely extends that for posets which we studied in the last section. (We should mention in passing that J. Koslowski [650] has proposed an alternative definition of 'continuous category', but it fails to have this 'conservativity' property.)

Proposition 4.2.2

(i) *For any locally small $\mathcal{C}$, the category* Ind-$\mathcal{C}$ *has filtered colimits.*

(ii) *An object of* Ind-$\mathcal{C}$ *is finitely presentable iff it is a retract of an object of the form $y(A)$, $A \in$* ob $\mathcal{C}$.

(iii) *If $\mathcal{C}$ is Cauchy-complete, then it is recoverable (up to equivalence) from* Ind-$\mathcal{C}$, *as the full subcategory* (Ind-$\mathcal{C})_\omega$ *of finitely presentable objects.*

(iv) *If $\mathcal{C}$ is cocartesian then so is* Ind-$\mathcal{C}$, *and the embedding* $y\colon \mathcal{C} \to \text{Ind-}\mathcal{C}$ *preserves finite colimits.*

(v) *If $\mathcal{C}$ is small and cocartesian, then* Ind-$\mathcal{C}$ *is complete and cocomplete.*

Proof (i) Let $(A_i \mid i \in I)$ be the vertices of a filtered diagram in Ind-$\mathcal{C}$, and suppose each A_i is the ind-object $(B_{ij} \mid j \in J_i)$. We define a small category K as follows: its objects are pairs (i,j) with $i \in I$ and $j \in J_i$, and morphisms $(i,j) \to (i',j')$ are pairs (α, f) where $\alpha\colon i \to i'$ in I and $f\colon B_{ij} \to B_{i'j'}$ is a representative for the jth component of the αth edge of the filtered diagram. Composition of morphisms of K is defined in the obvious way; we claim that it is a filtered category. We shall verify the third condition for filteredness (cf. B2.6.2(a)) – the other two are similar. Let

$$(i,j) \underset{(\alpha',f')}{\overset{(\alpha,f)}{\rightrightarrows}} (i',j')$$

be a parallel pair of morphisms in K. Since I is filtered, we can find $\beta\colon i' \to i''$ such that $\beta\alpha = \beta\alpha'$. Let $g\colon B_{i'j'} \to B_{i''j''}$ be any representative of the j'th component of the βth edge of the diagram: the composites gf and gf' need not be equal, but they both represent the jth component of the $\beta\alpha$th edge of the diagram, so they are coequalized by some $B_{i''j''} \to B_{i''j'''}$. Composing this morphism with g, we obtain a morphism g' such that (β, g') coequalizes (α, f) and (α', f') in K.

The assignment $(i,j) \mapsto B_{ij}$ becomes a functor $K \to \mathcal{C}$, and hence an object of Ind-$\mathcal{C}$, in an obvious way. Moreover, we have a morphism of ind-objects $(B_{ij} \mid j \in J_i) \to (B_{ij} \mid (i,j) \in K)$ for each i, whose jth component is represented by the identity morphism on B_{ij}. It is straightforward to verify that these morphisms form a cone under the original filtered diagram, and that it has the universal property of a colimit.

(ii) It follows easily from the construction of filtered colimits in (i) that they are preserved by the functors Ind-$\mathcal{C}\,(y(A), -)$. Thus any $y(A)$ is finitely presentable; hence so is any retract of such an object. Conversely, if $A = (A_i \mid i \in I)$ is finitely presentable in Ind-$\mathcal{C}$, then by expressing it as the colimit of the $y(A_i)$ we see that the identity morphism on A must factor through one of the maps $y(A_i) \to A$.

(iii) follows at once from (ii) and the fact that y is a full embedding.

(iv) It is clear that if 0 is initial in $\mathcal{C}$ then $y(0)$ is initial in Ind-$\mathcal{C}$. If $(A_i \mid i \in I)$ and $(B_j \mid j \in J)$ are two ind-objects, we define their coproduct to be the $I \times J$-indexed diagram whose (i,j)th vertex is $A_i + B_j$; the coprojection $(A_i \mid i \in I) \to (A_i + B_j \mid i \in I, j \in J)$ has as its ith component the equivalence class containing the coprojections $A_i \to A_i + B_j$ for all j, and it is straightforward to verify the universal property of the coproduct (and that $y(A+B) \cong y(A) + y(B)$). To form

the coequalizer of two morphisms

$$(A_i \mid i \in I) \underset{\psi}{\overset{\phi}{\rightrightarrows}} (B_j \mid j \in J),$$

we first form the category K whose objects are quadruples (i, j, f, g) where $i \in I$, $j \in J$ and $f, g \colon A_i \rightrightarrows B_j$ represent the ith components of ϕ and ψ respectively, and whose morphisms $(i, j, f, g) \to (i', j', f', g')$ are pairs $(\alpha \colon i \to i', \beta \colon j \to J')$ such that the squares

$$\begin{tikzcd}
A_i \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] \arrow[d, "A_\alpha"] & B_j \arrow[d, "B_\beta"] \\
A_{i'} \arrow[r, shift left, "f'"] \arrow[r, shift right, "g'"'] & B_{j'}
\end{tikzcd}$$

both commute; for $k = (i, j, f, g) \in K$, we define C_k to be the coequalizer of f and g in $\mathcal{C}$. Then it is straightforward to verify that K is filtered, and that the C_k define an ind-object $(C_k \mid k \in K)$. Not every object j of J necessarily occurs as the second component of an object of K, but for each j we can find a morphism $j \to j'$ such that j' occurs in an object (i, j', f, g), and thus we have a composite morphism $B_j \to B_{j'} \to C_{(i,j',f,g)}$; all such morphisms are moreover equivalent as morphisms into the ind-object $(C_k \mid k \in K)$, and we take their equivalence class to be the jth component of a morphism $(B_j \mid j \in J) \to (C_k \mid k \in K)$. This morphism is readily seen to have the universal property of a coequalizer of ϕ and ψ. And it is again easy to see that the embedding y preserves coequalizers.

(v) Cocompleteness follows from (i) and (iv). For completeness, it is easier to use the alternative representation of Ind-$\mathcal{C}$ as the closure of the full subcategory of representables under filtered colimits in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$; for in this case we may identify the latter with the category $\mathfrak{Cart}(\mathcal{C}^{\mathrm{op}}, \mathbf{Set})$ of cartesian functors, by B3.2.5, and this is clearly closed under limits in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. (Alternatively, we could use the Adjoint Functor Theorem to construct right adjoints for the diagonal functors Ind-$\mathcal{C} \to [\mathcal{D}, \text{Ind-}\mathcal{C}]$ for arbitrary small categories $\mathcal{D}$: note that the objects in the image of y form a separating set for Ind-$\mathcal{C}$.) □

Of course, the categories Ind-$\mathcal{C}$ where $\mathcal{C}$ is small (resp. small and cocartesian) are exactly the finitely accessible (resp. locally finitely presentable) categories studied in Section D2.3.

We digress for a moment to fulfil a promise made in Section C2.2. First we need:

Lemma 4.2.3 *Let $\mathcal{C}$ be a small cocartesian category with pullbacks, such that finite colimits are stable under pullback in $\mathcal{C}$. Then* Ind-$\mathcal{C}$ *is locally cartesian closed.*

Proof We use the identification of Ind-$\mathcal{C}$ with $\mathfrak{Cart}(\mathcal{C}^{\text{op}}, \textbf{Set})$. We recall from A1.1.7 that, for any functor $F: \mathcal{C}^{\text{op}} \to \textbf{Set}$, the slice category $[\mathcal{C}^{\text{op}}, \textbf{Set}]/F$ may be identified with $[\mathcal{F}^{\text{op}}, \textbf{Set}]$, where $\mathcal{F} \to \mathcal{C}$ is the discrete fibration corresponding to F. If F is cartesian, then it is easily verified that $\mathcal{F}$ inherits finite colimits and pullbacks, and the stability of the former under the latter, from $\mathcal{C}$. Moreover, given a natural transformation $G \to F$ between functors $\mathcal{C}^{\text{op}} \rightrightarrows \textbf{Set}$, the functor G is cartesian iff the corresponding functor $\mathcal{F}^{\text{op}} \to \textbf{Set}$ is cartesian; so we may identify $\mathfrak{Cart}(\mathcal{C}^{\text{op}}, \textbf{Set})/F$ with $\mathfrak{Cart}(\mathcal{F}^{\text{op}}, \textbf{Set})$. Thus it suffices to show that $\mathfrak{Cart}(\mathcal{C}^{\text{op}}, \textbf{Set})$ is cartesian closed; we shall do this by showing that it is an exponential ideal in $[\mathcal{C}^{\text{op}}, \textbf{Set}]$.

Suppose F is a cartesian functor $\mathcal{C}^{\text{op}} \to \textbf{Set}$, and let G be an arbitrary functor; then by the proof of A1.5.5 we know that the exponential F^G in $[\mathcal{C}^{\text{op}}, \textbf{Set}]$ may be defined by taking $F^G(A)$ to be the set of natural transformations $h(A) \times G \to F$, for each object A of $\mathcal{C}$. Now suppose A is the colimit of a finite diagram with vertices A_i, and suppose we are given a compatible family of natural transformations $\theta_i: h(A_i) \times G \to F$. We may construct $\theta: h(A) \times G \to F$ as follows: given an element $x \in G(B)$ and a morphism $B \to A$, we pull back the morphisms $A_i \to A$ along the latter to obtain a colimit diagram $B \cong \varinjlim B_i$, and then restrict x to a family of elements $x_i \in G(B_i)$. Then we obtain a family of elements $\theta_i(x_i, B_i \to A_i) \in F(A_i)$; but since F is cartesian these correspond to a unique element of $F(A)$, which we define to be $\theta(x, B \to A)$. It is straightforward to verify that this does define a natural transformation θ as required, and that it sets up a natural bijection between $F^G(A)$ and $\varprojlim F^G(A_i)$. □

Example 4.2.4 We may now give our promised example (due to F. Borceux and M. C. Pedicchio [156]) of a locally cartesian closed, locally finitely presentable category which is not a quasitopos. Let $\mathcal{C}$ be the category $\textbf{Set}_c$ of finite or countably infinite sets: $\mathcal{C}$ clearly satisfies the hypotheses of 4.2.3, and so Ind-$\mathcal{C}$ is locally cartesian closed. We claim next that Ind-$\mathcal{C}$ is balanced (so that, by A2.6.3(ii), it is a quasitopos iff it is a topos); for this we need only the fact that every epimorphism in $\mathcal{C}$ is regular. For if $A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ is both monic and epic in Ind-$\mathcal{C}$, then so is its pullback along any $y(C) \to B$, and since B is a colimit of objects in the image of y (and this colimit is stable under pullback along $A \rightarrowtail\!\!\!\!\twoheadrightarrow B$) it suffices to show that these pullbacks are all isomorphisms. So consider a morphism $\phi: (A_i \mid i \in I) \rightarrowtail\!\!\!\!\twoheadrightarrow y(C)$ which is both monic and epic. By the construction of 4.2.2(iv), we may identify the cokernel-pair of ϕ with the ind-object $(C +_{A_i} C \mid i \in I)$ whose ith vertex is the cokernel-pair of the ith component of ϕ; so this ind-object must be isomorphic to $y(C)$, and since the latter is finitely presentable it follows that the coprojections $C \rightrightarrows C +_{A_i} C$ must be equal for some i, i.e. some $A_i \to C$ is (regular) epic. But then $y(A_i) \to y(C)$ is regular epic in Ind-$\mathcal{C}$, since y preserves coequalizers; and since it factors through the monomorphism ϕ, the latter must be an isomorphism.

Now we observe that the set $\mathbb{N}$ of natural numbers satisfies Freyd's axioms for a natural number object (A2.5.5) in $\mathcal{C}$, so by 4.2.2(iv) $y(\mathbb{N})$ must satisfy them

in Ind-$\mathcal{C}$, and if the latter were a topos then $y(\mathbb{N})$ would be its natural number object. But $y(\mathbb{N})$ is finitely presentable in Ind-$\mathcal{C}$; and the natural number object cannot be finitely presentable in any non-degenerate cocomplete topos, since it is a countable copower of 1 and thus expressible as the filtered colimit of the finite copowers of 1. So Ind-$\mathcal{C}$ is not a topos; hence it is not a quasitopos, either.

The above counterexample should be compared with D3.3.9. In the latter example, the codomain of the Yoneda embedding is a topos, but the embedding does not preserve coequalizers; here, the embedding preserves coequalizers, but its codomain is not a topos.

We now return to the general study of ind-completions. The mapping $\mathcal{C} \mapsto$ Ind-$\mathcal{C}$ is functorial, and indeed 2-functorial, on $\mathfrak{CAT}$, in an obvious way; and the embedding $y\colon \mathcal{C} \to$ Ind-$\mathcal{C}$ is the unit of a monad structure on this 2-functor. In fact this monad is a KZ-monad, as defined in B1.1.11; so an algebra structure for it on a category $\mathcal{C}$ is the same thing as a left adjoint for y. Explicitly, we have

Lemma 4.2.5 *The embedding $y\colon \mathcal{C} \to$ Ind-$\mathcal{C}$ has a left adjoint iff $\mathcal{C}$ has filtered colimits.*

Proof Suppose the left adjoint exists; let us call it F for the moment. For any object $(A_i \mid i \in I)$ of Ind-$\mathcal{C}$ and any object B of $\mathcal{C}$, we have

$$\mathcal{C}\,(F(A_i \mid i \in I), B) \cong \text{Ind-}\mathcal{C}\,((A_i \mid i \in I), y(B)) = \varprojlim_{i \in I} \mathcal{C}\,(A_i, B),$$

so $F(A_i \mid i \in I)$ must be an actual colimit in $\mathcal{C}$ of this filtered diagram. Conversely, if $\mathcal{C}$ has filtered colimits, then by defining $F(A_i \mid i \in I)$ to be such a colimit we clearly obtain a functor Ind-$\mathcal{C} \to \mathcal{C}$ left adjoint to y. $\square$

In view of the above result, we denote the left adjoint of y when it exists by $\varinjlim$.

Corollary 4.2.6 *Ind-$\mathcal{C}$ is the free filtered-colimit completion of $\mathcal{C}$; that is, for any category $\mathcal{D}$ with filtered colimits, any functor $F\colon \mathcal{C} \to \mathcal{D}$ extends, uniquely up to canonical isomorphism, to a filtered-colimit-preserving functor* Ind-$\mathcal{C} \to \mathcal{D}$.

Proof Uniqueness of the extension is immediate from the fact that every object of Ind-$\mathcal{C}$ is a filtered colimit of objects of $\mathcal{C}$. For the existence, consider the composite

$$\text{Ind-}\mathcal{C} \xrightarrow{\text{Ind-}F} \text{Ind-}\mathcal{D} \xrightarrow{\varinjlim} \mathcal{D};$$

note that a functor of the form Ind-F always preserves filtered colimits, from the way in which we constructed them in 4.2.2. $\square$

Definition 4.2.7 We define a category $\mathcal{C}$ to be *continuous* if it has filtered colimits and the functor $\varinjlim$: Ind-$\mathcal{C} \to \mathcal{C}$ itself has a left adjoint. We shall denote this left adjoint by w if it exists.

Paralleling 4.1.4, we have what is essentially a special case of B1.1.15:

Proposition 4.2.8 *For a category $\mathcal{C}$ with filtered colimits, the following are equivalent:*

(i) *$\mathcal{C}$ is continuous.*

(ii) *$\mathcal{C}$ is a coadjoint retract of a category of the form* Ind-$\mathcal{D}$ *by functors preserving filtered colimits.*

(iii) *$\mathcal{C}$ is a retract of a category of the form* Ind-$\mathcal{D}$ *by functors preserving filtered colimits.*

(iv) *$\mathcal{C}$ is a retract of a continuous category by functors preserving filtered colimits.*

Proof (i) $\Rightarrow$ (ii) is trivial: note that the unit of the adjunction ($w \dashv \varinjlim$) is necessarily an isomorphism, since the counit of ($\varinjlim \dashv y$) is an isomorphism.

(ii) $\Rightarrow$ (iii) is trivial.

(iii) $\Rightarrow$ (iv): We have to show that categories of the form Ind-$\mathcal{D}$ are continuous; we have already observed that they have filtered colimits and that functors of the form Ind-F preserve them, so that $y\colon$ Ind-$\mathcal{D} \to$ Ind-Ind-$\mathcal{D}$ is right adjoint to $\varinjlim$: Ind-Ind-$\mathcal{D} \to$ Ind-$\mathcal{D}$, naturally in $\mathcal{D}$. It now follows from the general theory of KZ-monads (see B1.1.12) that Ind-$y\colon$ Ind-$\mathcal{D} \to$ Ind-Ind-$\mathcal{D}$ is left adjoint to $\varinjlim$. In fact it is not hard to show this directly: let $(B_i \mid i \in I)$ be a filtered diagram in Ind-$\mathcal{D}$, and suppose that $B = \varinjlim(B_i \mid i \in I)$ is the filtered diagram $(A_j \mid j \in J)$ in $\mathcal{D}$. Then, because the $y(A_j)$ are finitely presentable in Ind-$\mathcal{D}$, the canonical morphisms $y(A_j) \to B$ in Ind-$\mathcal{D}$ each factor in an essentially unique way through some leg $B_i \to B$ of the colimiting cone. In this way we obtain a well-defined morphism

$$\text{Ind-}y(B) = (y(A_j) \mid j \in J) \longrightarrow (B_i \mid i \in I)$$

in Ind-Ind-$\mathcal{D}$, whose image under $\varinjlim$ is the identity map on B. It is now straightforward to verify that this morphism is a component of a natural transformation from Ind-$y \cdot \varinjlim$ to the identity functor on Ind-Ind-$\mathcal{D}$, which satisfies the triangular identities with the isomorphism $1_{\text{Ind-}\mathcal{D}} \cong \varinjlim \cdot \text{Ind-}y$. So we have the required adjunction.

(iv) $\Rightarrow$ (i): Suppose $\mathcal{C}$ is a retract of a continuous category $\mathcal{D}$. Then we may apply B1.1.10 to the diagram

$$\begin{tikzcd}
\text{Ind-}\mathcal{C} \arrow[r, "\text{Ind-}i"] \arrow[d, "\varinjlim"] & \text{Ind-}\mathcal{D} \arrow[r, "\text{Ind-}r"] \arrow[d, "\varinjlim"] & \text{Ind-}\mathcal{C} \arrow[d, "\varinjlim"] \\
\mathcal{C} \arrow[r, "i"] & \mathcal{D} \arrow[r, "r"] & \mathcal{C}
\end{tikzcd}$$

whose commutativity is just the statement that r and i preserve filtered colimits. (To do this, we need to consider the diagram above as lying in a (meta-)2-category in which idempotent 2-cells split; $\mathfrak{CAT}$ will not serve for this purpose, but in fact the diagram lies in the full sub-2-category $\mathfrak{CAU}$ of Cauchy-complete categories (cf. A1.1.10), since all the categories in it have filtered colimits. Recall that splitting an idempotent may be regarded as forming the colimit of a diagram over the two-element monoid $M = \{1, e\}$ with $e^2 = e$, and this monoid is a filtered category.) □

The next two results are sometimes helpful in identifying which categories are continuous.

Lemma 4.2.9 *If $\mathcal{C}$ is a continuous category, then so is $\mathcal{C}/A$ for any object A of $\mathcal{C}$.*

Proof From the 'double-limit' definition of morphisms in Ind-$\mathcal{C}$, it is easy to see that a morphism from an arbitrary ind-object $(B_i \mid i \in I)$ to $y(A)$ is the same thing as a cone under $(B_i \mid i \in I)$ with vertex A, and hence that Ind-$(\mathcal{C}/A)$ is isomorphic to Ind-$\mathcal{C}/y(A)$. It is now easy to see that if $\mathcal{C}$ is continuous then we may 'lift' the left adjoint w of $\varinjlim : \text{Ind-}\mathcal{C} \to \mathcal{C}$ to a left adjoint for $\varinjlim : \text{Ind-}\mathcal{C}/y(A) \to \mathcal{C}/A$. □

Proposition 4.2.10 *Let $\mathcal{C}$ be a category with filtered colimits and pullbacks. Then $\mathcal{C}$ is continuous iff*

(i) *filtered colimits are stable under pullback in $\mathcal{C}$, and*

(ii) *for each object A of $\mathcal{C}$, the category Ind-$\mathcal{C}_A$ of ind-objects in $\mathcal{C}$ with colimit A, and morphisms between them inducing the identity morphism on A, has an initial object.*

Proof First we show that if $\mathcal{C}$ is continuous and has pullbacks then it necessarily satisfies (i). (This is, of course, the analogue for categories of the first assertion of 4.1.2, that a continuous semilattice is a preframe.) Suppose given an ind-object $(B_i \mid i \in I)$, and a morphism $f : A \to \varinjlim_I B_i = B$. Then we may make the pullbacks $A \times_B B_i$ into the vertices of an I-indexed diagram, with colimit A' say, and we clearly have an induced morphism $g : A' \to A$. But we may transpose f to a morphism $w(A) \to (B_i \mid i \in I)$ in Ind-$\mathcal{C}$; and each representative $A_j \to B_i$ for the jth component of this morphism may be factored through the corresponding $A \times_B B_i$, yielding a morphism of ind-objects $w(A) \to (A \times_B B_i \mid i \in I)$. Applying the functor $\varinjlim$ to this morphism yields a morphism $h : A \to A'$ in $\mathcal{C}$. Now the composite $gh : A \to A' \to A$ is the identity, since we may regard it as the effect of applying $\varinjlim$ to the composite $w(A) \to (A \times_B B_i \mid i \in I) \to y(A)$; hence in particular h is split monic. But since $fgh = f$, every morphism in the colimit cone $(A \times_B B_i \to A' \mid i \in I)$ factors through h; so g and h are inverse isomorphisms.

Given that (i) holds, it is now easy to see that the functor $\varinjlim : \text{Ind-}\mathcal{C} \to \mathcal{C}$ is a fibration in the sense of B1.3.4: to obtain a prone lifting of a morphism $A \to B = \varinjlim(B_i \mid i \in I)$ in $\mathcal{C}$, we form the ind-object $(A \times_B B_i \mid i \in I)$, and the morphism from this ind-object to $(B_i \mid i \in I)$ whose ith component is the equivalence class of the projection $A \times_B B_i \to B_i$. And it is well known that specifying an initial object in each of the fibres of a fibration $\Pi : \mathcal{S} \to \mathcal{T}$ is equivalent to specifying a left adjoint for Π such that the unit of the adjunction is an isomorphism. □

Corollary 4.2.11 *Let $\mathcal{C}$ be a continuous category with pullbacks and images. Then, for any object A of $\mathcal{C}$, $\mathrm{Sub}_{\mathcal{C}}(A)$ is a continuous semilattice.*

Proof By 4.2.9, it suffices to consider the case (when $\mathcal{C}$ has a terminal object and) $A = 1$. In this case, we know that the support functor $\sigma : \mathcal{C} \to \mathrm{Sub}_{\mathcal{C}}(1)$ preserves filtered colimits, since it is left adjoint to the inclusion; we claim that the inclusion does so too (i.e. that a filtered colimit in $\mathcal{C}$ of subterminal objects is subterminal), from which the result will follow by 4.2.8 (plus 4.2.1). But if $(U_i \mid i \in I)$ is a filtered diagram of subterminal objects, we have

$$\varinjlim_I U_i \times \varinjlim_I U_i \cong \varinjlim_{I \times I}(U_i \times U_j)$$

by two applications of condition (i) of 4.2.10; and the latter is isomorphic to $\varinjlim_I(U_i \times U_i)$ since the diagonal $I \to I \times I$ is a final functor. So $\varinjlim_I U_i$ is subterminal. □

Next, we introduce the analogue for continuous categories of the way-below relation which we used in defining the notion of continuous poset. We note that, in a continuous poset B, the relation $a << b$ holds iff $\downarrow(a) \subseteq \downarrow\!\!(b)$ in IB; this provides the motivation for the following definition.

Definition 4.2.12 Let $\mathcal{C}$ be a continuous category, A and B two objects of $\mathcal{C}$. By a *wavy arrow* from A to B in $\mathcal{C}$, we mean a morphism $y(A) \to w(B)$ in Ind-$\mathcal{C}$. We write $f : A \rightsquigarrow B$ for 'f is a wavy arrow from A to B'.

Equivalently, if $w(B)$ is the ind-object $(B_i \mid i \in I)$, then a wavy arrow $A \rightsquigarrow B$ is an equivalence class of 'straight' arrows $A \to B_i$ in $\mathcal{C}$, where the equivalence relation identifies $(A \to B_i)$ with $(A \to B_{i'})$ iff there exist $i \to i''$ and $i' \to i''$ in I such that the diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B_i \arrow[d] \\ B_{i'} \arrow[r] & B_{i''} \end{tikzcd}$$

commutes.

Since $\varinjlim y(A) \cong A$ and $\varinjlim w(B) \cong B$, applying $\varinjlim$ to a wavy arrow $f\colon A \rightsquigarrow B$ yields an *underlying straight arrow* $A \to B$, which we denote by $\epsilon(f)$. However, the mapping $f \mapsto \epsilon(f)$ is not injective in general; thus we have to think of waviness as an additional structure carried by a morphism of $\mathcal{C}$, rather than as a property of such morphisms.

Example 4.2.13 Let G be a group, and consider the category $[G, \mathbf{Set}]$ of left G-sets. As in any locally finitely presentable category (cf. the proof of (iii)$\Rightarrow$(iv) in 4.2.8), the ind-object $w(B)$, for a G-set B, may be identified with the filtered diagram of all morphisms from finitely presentable G-sets to B; in particular, $w(1)$ is the filtered diagram of all finitely presentable G-sets. And it is easy to see that a G-set is finitely presentable iff it has finitely many orbits, and the stabilizer of each element is a finitely generated subgroup of G (cf. D2.4.1). In particular, if G itself is not finitely generated, then the projections $G \times G \rightrightarrows G$ both define wavy arrows $G \times G \rightsquigarrow 1$ in $[G, \mathbf{Set}]$, which are not equal since the projections cannot be coequalized by any map from G to a finitely presentable G-set. But they clearly have the same underlying straight arrow.

A wavy arrow $f\colon A \rightsquigarrow B$ may be composed with a straight arrow $g\colon A' \to A$ or $h\colon B \to B'$, by forming $f \cdot y(g)$ or $w(h) \cdot f$ as appropriate; these compositions are associative and commute with each other, so that the function $\mathcal{W}$ sending (A, B) to the set of all wavy arrows $A \rightsquigarrow B$ may be viewed as a profunctor $\mathcal{C} \nrightarrow \mathcal{C}$, as defined in Section B2.7. In addition, we have

Lemma 4.2.14 *Let $f\colon A \rightsquigarrow B$ and $g\colon B \rightsquigarrow C$ be two wavy arrows in a continuous category $\mathcal{C}$. Then the composites $g \cdot \epsilon(f)$ and $\epsilon(g) \cdot f$ are equal as wavy arrows $A \rightsquigarrow C$.*

Proof It is easy to see that each is the composite

$$y(A) \xrightarrow{\ f\ } w(B) \xrightarrow{\ i\ } y(B) \xrightarrow{\ g\ } w(C)$$

in Ind-$\mathcal{C}$. where i is the transpose across $(w \dashv \varinjlim)$ of the isomorphism $B \cong \varinjlim y(B)$ (equivalently, the transpose across $(\varinjlim \dashv y)$ of $\varinjlim w(B) \cong B$). $\square$

Lemma 4.2.14 tells us that we have a well-defined composition for pairs of wavy arrows; and it follows easily from the definitions that the eight possible associative laws $h \cdot (g \cdot f) = (h \cdot g) \cdot f$, where each of f, g and h may be either wavy or straight, are all satisfied. In particular, taking g to be straight and the other two to be wavy, we should be entitled to regard composition of wavy arrows as a morphism of profunctors $\mu\colon \mathcal{W} \otimes_{\mathcal{C}} \mathcal{W} \to \mathcal{W}$, if only the domain of this morphism were defined. The problem is that, in order to form $(\mathcal{W} \otimes_{\mathcal{C}} \mathcal{W})(A, B)$, we have to form a suitable quotient of the disjoint union $\coprod_{C \in \mathrm{ob}\, \mathcal{C}} \mathcal{W}(C, B) \times \mathcal{W}(A, C)$: and since $\mathcal{C}$ is not a small category, the latter will be a proper class in general.

However, we note that each wavy arrow $C \rightsquigarrow B$ can be factored as $C \to B_i \rightsquigarrow B$, where B_i is one of the vertices of the ind-object $w(B) = (B_i \mid i \in I)$, and the wavy arrow $B_i \rightsquigarrow B$ corresponds to the identity morphism on B_i. Since

there is only a set of such B_i, for each B, it follows that we may reduce the proper class above to the set $\coprod_{i\in I} \mathcal{W}(B_i, B) \times \mathcal{W}(A, B_i)$; more explicitly, we may define $(\mathcal{W} \otimes_{\mathcal{C}} \mathcal{W})(A, B)$ to be $\varinjlim_{i\in I} \mathcal{W}(A, B_i)$. Thus μ becomes a well-defined morphism of profunctors.

The following result is the analogue for continuous categories of the interpolation property for the way-below relation in a continuous poset.

Proposition 4.2.15 *The morphism $\mu\colon \mathcal{W} \otimes_{\mathcal{C}} \mathcal{W} \to \mathcal{W}$ just defined is an isomorphism of profunctors.*

Proof To show that μ is surjective, we must show that each wavy arrow can be factored as a composite of two wavy arrows. But if $w(A)$ is the ind-object $(A_i \mid i \in I)$, and $w(A_i) = (B_{ij} \mid j \in J_i)$ for each i, then the $w(A_i)$ form a diagram of shape I in Ind-$\mathcal{C}$, and we can form its colimit $(B_{ij} \mid (i,j) \in K)$ as in 4.2.2(i). It is clear that a cone under this filtered diagram in $\mathcal{C}$ must factor uniquely through each A_i and hence through A; that is, $\varinjlim (B_{ij} \mid (i,j) \in K) \cong A$. Transposing, we get a morphism of ind-objects $w(A) \to (B_{ij} \mid (i,j) \in K)$; hence every $A_i \to A$ factors through some $B_{i'j} \to A$. This suffices to prove surjectivity.

For injectivity, suppose we have a commutative square of wavy arrows

$$\begin{tikzcd} A \arrow[r, squiggly, "f"] \arrow[d, squiggly, "g"'] & B \arrow[d, squiggly, "h"] \\ C \arrow[r, squiggly, "k"] & D \end{tikzcd} .$$

Since $w(D)$ is filtered, we may respresent h and k by morphisms $h_i\colon B \to D_i$ and $k_i\colon C \to D_i$ into the same vertex D_i of $w(D)$. Now the square

$$\begin{tikzcd} A \arrow[r, "\epsilon(f)"] \arrow[d, "\epsilon(g)"] & B \arrow[d, "h_i"] \\ C \arrow[r, "k_i"] & D_i \end{tikzcd}$$

need not commute, but both ways round represent the same wavy arrow $A \rightsquigarrow D$, so we can find $D_i \to D_{i'}$ in $w(D)$ coequalizing them. Even so, the square

$$\begin{tikzcd} A \arrow[r, squiggly, "f"] \arrow[d, squiggly, "g"'] & B \arrow[d, "h_{i'}"] \\ C \arrow[r, "k_{i'}"] & D_{i'} \end{tikzcd}$$

need not commute at the wavy level; but since the two ways round it have the same underlying straight arrow, it follows from 4.2.14 that they have the same composite with (the underlying straight arrow of) any wavy arrow with domain $D_{i'}$. So we may use the first part of the proof to factor the wavy arrow $D_{i'} \rightsquigarrow D$ as a composite $D_{i'} \rightsquigarrow D_{i''} \rightsquigarrow D$, and replace $h_{i'}$ and $k_{i'}$ by their composites with the underlying straight arrow of $D_{i'} \rightsquigarrow D_{i''}$. We thus have diagrams

$$\begin{tikzcd} A \arrow[r, squiggly] \arrow[d, squiggly, "f"'] & D_{i''} \arrow[d, squiggly] \\ B \arrow[ur, "h_{i''}"] \arrow[r, squiggly, "h"] & D \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} A \arrow[r, squiggly] \arrow[d, squiggly, "g"'] & D_{i''} \arrow[d, squiggly] \\ C \arrow[ur, "k_{i''}"] \arrow[r, squiggly, "k"] & D \end{tikzcd}$$

in which all cells commute at the wavy level, which shows that (h, f) and (k, g) are equal as elements of $\mathcal{W} \otimes \mathcal{W}(A, D)$. $\square$

In view of 4.2.15, we may consider the inverse of μ as a morphism of profunctors $\mathcal{W} \to \mathcal{W} \otimes \mathcal{W}$. Since μ is associative, μ^{-1} is coassociative; and from the way in which μ was defined it is easy to see that $\epsilon \colon \mathcal{W} \to \mathcal{C}(-,-)$ is a counit for μ^{-1} (recall from B2.7.3 that the unit for profunctor composition is the Yoneda profunctor $\mathcal{C}(-,-)$). We thus have

Theorem 4.2.16 *For any continuous category $\mathcal{C}$, the structure $(\mathcal{W}, \epsilon, \mu^{-1})$ is an idempotent profunctor comonad on $\mathcal{C}$.* $\square$

Of course, 4.2.16 has to be read *cum grano salis*, since profunctor composition for profunctors between large categories is not defined in general (as we saw, it is merely a happy accident that $\mathcal{W} \otimes \mathcal{W}$ is defined), and so we do not have a genuine bicategory for this comonad to live in. However, 4.2.19 below provides a more legitimate version of the theorem, applicable to all the continuous categories we shall be interested in.

A further property of wavy arrows which we shall require is the following:

Lemma 4.2.17 *Let $\mathcal{C}$ be a continuous category. Then any finite colimits which exist in $\mathcal{C}$ are also 'wavy colimits'; that is, given a finite (straight-edged) diagram in $\mathcal{C}$ with vertices $(A_1, \ldots, A_n)$ and a colimit cone $(\lambda_i \colon A_i \to L \mid 1 \leq i \leq n)$, if we are given any 'wavy cone' under the diagram (that is, a family of wavy arrows $(f_i \colon A_i \rightsquigarrow B \mid 1 \leq i \leq n)$ which commute with all the edges of the diagram), then there exists a unique wavy arrow $f \colon L \rightsquigarrow B$ such that $f\lambda_i = f_i$ for all i.*

Proof Suppose given a wavy cone as above. Let $(B_j \mid j \in J)$ be the vertices of the ind-object $w(B)$; then we can represent each f_i by a straight arrow $g_i \colon A_i \to B_j$ for some j, and since there are finitely many of them we can choose the same j for each i. Even so, given an edge $A_i \to A_{i'}$ of the diagram, the

triangle

$$\begin{tikzcd} A_i \arrow[r] \arrow[dr, "g_i"'] & A_{i'} \arrow[d, "g_{i'}"] \\ & B_j \end{tikzcd}$$

need not commute; but the two ways round it are coequalized by some $B_j \to B_{j'}$, since they represent the same wavy arrow $A_i \rightsquigarrow B$. Again, since there are finitely many edges in the diagram, we can choose a single $B_j \to B_{j'}$ coequalizing all such triangles; then we have a straight-edged cone under the diagram with vertex $B_{j'}$, and hence an induced morphism $g: L \to B_{j'}$, which we can regard as representing a wavy arrow $f: L \rightsquigarrow B$.

This establishes the existence of f; for the uniqueness, suppose we had another such morphism f'. We could represent f' by a straight arrow $g': L \to B_{j''}$ for some j''; then, for each i, the composite $g'\lambda_i$ would represent the same wavy arrow $A_i \rightsquigarrow B$ as g_i, and so we could find morphisms $j' \to j'''$ and $j'' \to j'''$ in J such that

$$\begin{tikzcd} A_i \arrow[r, "g_i"] \arrow[d, "g'\lambda_i"] & B_{j'} \arrow[d] \\ B_{j''} \arrow[r] & B_{j'''} \end{tikzcd}$$

commutes. Doing this for all the A_i, we obtain a vertex $B_{j'''}$ of $w(B)$ such that, when we compose g and g' with the morphisms into $B_{j'''}$, they become factorizations of the same cone through L, and therefore equal; so f and f' are equal as wavy arrows $L \rightsquigarrow B$. □

We shall be interested in two 'refinements' of 4.2.8, dealing with the cases when $\mathcal{C}$ is a retract of Ind-$\mathcal{D}$ for some *small* category $\mathcal{D}$, and when $\mathcal{C}$ is cocomplete. In connection with the first of these, we note that if $\mathcal{C} \cong$ Ind-$\mathcal{D}$ where $\mathcal{D}$ is small, then $\mathcal{C}$ contains a small full subcategory such that every object of $\mathcal{C}$ is expressible as a filtered colimit of objects of the subcategory. We shall say that a category $\mathcal{C}$ with filtered colimits is *ind-small* if it contains a small full subcategory $\mathcal{D}$ with this property; more precisely, if there exists a functor $h: \mathcal{C} \to$ Ind-$\mathcal{D}$ such that the composite

$$\mathcal{C} \xrightarrow{h} \text{Ind-}\mathcal{D} \longrightarrow \text{Ind-}\mathcal{C} \xrightarrow{\varinjlim} \mathcal{C}$$

is isomorphic to the identity. A sufficient condition for this is that $\mathcal{D}$ should be a *dense* full subcategory (i.e. one such that, for each $A \in \text{ob}\,\mathcal{C}$, the canonical cone under the forgetful diagram $(\mathcal{D}\downarrow A) \to \mathcal{C}$ is colimiting) for which the comma

category $(\mathcal{D} \downarrow A)$ is filtered for all A. (If $\mathcal{C}$ is cocartesian, the latter condition will hold provided $\mathcal{D}$ is closed under finite colimits in $\mathcal{C}$.) However, we do not actually require $\mathcal{D}$ to be dense – though it is so when (as in the case $\mathcal{C} = \text{Ind-}\mathcal{D}$) its objects are finitely presentable in $\mathcal{C}$, since then the obvious functor from (the index category of) $h(A)$ to $(\mathcal{D} \downarrow A)$ is final.

Proposition 4.2.18 *For a locally small category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ is continuous and ind-small.*

(ii) *$\mathcal{C}$ expressible as a retract, by filtered-colimit-preserving functors, of a category of the form* Ind-$\mathcal{D}$ *where $\mathcal{D}$ is small (equivalently, of a finitely accessible category).*

(iii) *$\mathcal{C}$ is expressible as a coadjoint retract of a category as in* (ii).

Proof (iii) $\Rightarrow$ (ii) is obvious. For (ii) $\Rightarrow$ (i), we have already observed that a category of the form Ind-$\mathcal{D}$ with $\mathcal{D}$ small is (continuous and) ind-small, so we need only show that this property is inherited by retracts. So suppose we have filtered-colimit-preserving functors

$$\mathcal{C} \xrightarrow{\quad i \quad} \tilde{\mathcal{C}} \xrightarrow{\quad r \quad} \mathcal{C}$$

with $ri \cong 1$, and a small subcategory $\tilde{\mathcal{D}}$ of $\tilde{\mathcal{C}}$ with the required property. Let $\mathcal{D}$ be the full subcategory of $\mathcal{C}$ on the objects $r(B)$, $B \in \text{ob}\,\tilde{\mathcal{D}}$. Then the composite

$$\mathcal{C} \xrightarrow{\quad i \quad} \tilde{\mathcal{C}} \xrightarrow{\quad \tilde{h} \quad} \text{Ind-}\tilde{\mathcal{D}} \xrightarrow{\text{Ind-}r} \text{Ind-}\mathcal{D}$$

is easily seen to have the required property, using the fact that r preserves filtered colimits.

(i) $\Rightarrow$ (iii): Suppose $\mathcal{C}$ is ind-small. Let $\mathcal{D}$ be the small full subcategory of $\mathcal{C}$ provided by the definition, and consider the composite

$$\mathcal{C} \xrightarrow{\quad w \quad} \text{Ind-}\mathcal{C} \xrightarrow{\text{Ind-}h} \text{Ind-Ind-}\mathcal{D} \xrightarrow{\varinjlim} \text{Ind-}\mathcal{D};$$

that is, given an object A of $\mathcal{C}$, if $w(A)$ is the ind-object $(A_i \mid i \in I)$ and $h(A_i) = (B_{ij} \mid j \in J_i)$ for each i, we consider the colimit of the (B_{ij}), constructed as in 4.2.2(i), as an ind-object $k(A)$ in $\mathcal{D}$. For each pair (i, j), we have a morphism $B_{ij} \to A_i$, and these morphisms clearly define a morphism $k(A) \to w(A)$ in Ind-$\mathcal{C}$; but it is also easy to see that we have $\varinjlim k(A) \cong A$, so the adjunction yields a morphism $w(A) \to k(A)$ in Ind-$\mathcal{C}$. A further straightforward verification shows that these two morphisms are inverse to each other, so we have shown that (up to isomorphism) the functor $w\colon \mathcal{C} \to \text{Ind-}\mathcal{C}$ factors through the full embedding $\text{Ind-}\mathcal{D} \to \text{Ind-}\mathcal{C}$ induced by $\mathcal{D} \subseteq \mathcal{C}$. As such, it is clearly still left adjoint to the restriction of $\varinjlim : \text{Ind-}\mathcal{C} \to \mathcal{C}$ to Ind-$\mathcal{D}$, and so it preserves filtered colimits; but so does the restriction of $\varinjlim$, since the inclusion $\text{Ind-}\mathcal{D} \to \text{Ind-}\mathcal{C}$ preserves them. $\square$

Remark 4.2.19 In the situation of 4.2.18, we may restrict $\mathcal{W}$ to a profunctor $\mathcal{D} \nrightarrow \mathcal{D}$; and, given that we have seen that the functor $w: \mathcal{C} \to \text{Ind-}\mathcal{C}$ may be taken to factor through Ind-$\mathcal{D}$, it is clear that the proof of its idempotency which we gave in 4.2.14 is still valid. That is, we now have a (legitimate) profunctor comonad $(\mathcal{W}, \epsilon, \mu^{-1})$ on $\mathcal{D}$. We shall see in the next section that the continuous category $\mathcal{C}$ may be reconstructed, up to equivalence, from $\mathcal{D}$ together with this data.

Concerning the second refinement mentioned above, we know from 4.2.2(iv) that Ind-$\mathcal{C}$ is cocomplete if $\mathcal{C}$ is locally small and cocartesian. The converse holds if $\mathcal{C}$ is Cauchy-complete, since in any category $\mathcal{D}$ with filtered colimits the full subcategory $\mathcal{D}_\omega$ of finitely presentable objects is closed under whatever finite colimits exist in $\mathcal{D}$.

Proposition 4.2.20 *For a locally small category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ is continuous and cocomplete.*

(ii) *$\mathcal{C}$ is a retract, by functors preserving filtered colimits, of a category of the form* Ind-$\mathcal{D}$ *where $\mathcal{D}$ is cocartesian.*

(iii) *$\mathcal{C}$ is a coadjoint retract of a category as in* (ii).

Proof (i) $\Rightarrow$ (iii) since continuity of $\mathcal{C}$ implies that it is a coadjoint retract of Ind-$\mathcal{C}$. (iii) $\Rightarrow$ (ii) is trivial; and for (ii) $\Rightarrow$ (i) we have merely to observe that $\mathcal{C}$ inherits cocompleteness from Ind-$\mathcal{D}$ – which is, once again, an easy application of B1.1.10, given that $\mathcal{C}$ necessarily inherits filtered colimits from Ind-$\mathcal{D}$, and is therefore Cauchy-complete. $\square$

We may also combine 4.2.18 and 4.2.20, as follows:

Corollary 4.2.21 *For a locally small category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ is continuous, ind-small and cocomplete.*

(ii) *$\mathcal{C}$ is a retract, by functors preserving filtered colimits, of a category of the form* Ind-$\mathcal{D}$ *where $\mathcal{D}$ is small and cocartesian (equivalently, of a locally finitely presentable category).*

(iii) *$\mathcal{C}$ is a coadjoint retract of a category as in* (ii).

Proof (iii) $\Rightarrow$ (ii) is as usual trivial, and (ii) $\Rightarrow$ (i) follows immediately from the corresponding parts of 4.2.18 and 4.2.20. For (i) $\Rightarrow$ (iii), suppose $\mathcal{C}$ is ind-small and cocomplete; let $\mathcal{D}$ be a small full subcategory satisfying the condition for ind-smallness. Then the closure $\overline{\mathcal{D}}$ of $\mathcal{D}$ under finite colimits in $\mathcal{C}$ is still small, and it also satisfies the condition; so, working through the proof of 4.2.18(i) $\Rightarrow$ (iii), we deduce that $\mathcal{C}$ is a coadjoint retract of Ind-$\overline{\mathcal{D}}$. $\square$

Suggestion for further reading: Johnstone & Joyal [542].

C4.3 Injective toposes

In this section, we consider the notion of injectivity for toposes corresponding to that for locales (or sober spaces) which we met in 4.1.6. (Throughout this and the next section, we shall work over a base topos $\mathcal{S}$, which is assumed to have a natural number object; but we shall generally treat it notationally as if it were the classical category of sets.) First, we must pay attention to the fact that we are working in a 2-category, and therefore need to take some care in defining the notion of 'injective object'. If we are given a class $\mathcal{M}$ of 1-cells in a 2-category $\mathfrak{K}$, the simplest notion of 'injective object with respect to $\mathcal{M}$' would be an object A such that, for every diagram

$$\begin{tikzcd} B \arrow[r, "m"] \arrow[d, "f"] & C \\ A & \end{tikzcd}$$

with $m \in \mathcal{M}$, there exists a morphism $g\colon C \to A$ and a 2-isomorphism $gm \cong f$. We shall call this notion *weak injectivity* with respect to $\mathcal{M}$. However, it seems reasonable that the operation of 'extending along m' should be defined for 2-cells as well as 1-cells with codomain A, which leads to the following stronger notion: A is *strongly injective* with respect to $\mathcal{M}$ if, for every $m\colon B \to C$ in $\mathcal{M}$, we are given a functor $m^+ : \mathfrak{K}(B, A) \to \mathfrak{K}(C, A)$ and a natural isomorphism fitting into the triangle

$$\begin{tikzcd} \mathfrak{K}(B,A) \arrow[r, "m^+"] \arrow[dr, "1"'] & \mathfrak{K}(C,A) \arrow[d, "(-)\circ m"] \\ & \mathfrak{K}(B,A) \end{tikzcd} .$$

A still stronger notion is that of a complete injective, which is dual to the notion of cocomplete injective that we met in B1.1.16. Explicitly, we say that A is a *complete $\mathcal{M}$-injective* if, for every morphism $f\colon B \to C$ of $\mathfrak{K}$, the functor $\mathfrak{K}(C, A) \to \mathfrak{K}(B, A)$ obtained by composition with f has a right adjoint f^+, the counit of the adjunction being an isomorphism whenever $f \in \mathcal{M}$.

Fortunately, in the 2-category $\mathfrak{BTop}/\mathcal{S}$ of bounded toposes over a base $\mathcal{S}$ with natural number object, the three notions of injectivity coincide:

Theorem 4.3.1 *Let $\mathcal{S}$ be a topos with natural number object. For a bounded $\mathcal{S}$-topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is a retract in $\mathfrak{BTop}/\mathcal{S}$ of a functor category $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ where $\mathbb{C}$ is cartesian.*

(ii) *$\mathcal{E}$ is a complete injective (with respect to the class of inclusions) in $\mathfrak{BTop}/\mathcal{S}$.*

(iii) *$\mathcal{E}$ is strongly injective in $\mathfrak{BTop}/\mathcal{S}$.*

(iv) *$\mathcal{E}$ is weakly injective in $\mathfrak{BTop}/\mathcal{S}$.*

Proof The implications (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) are clear. (iv) $\Rightarrow$ (i) follows from the fact that every bounded $\mathcal{S}$-topos $\mathcal{E}$ is a subtopos of $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ for some cartesian $\mathbb{C}$ – if $\mathcal{E}$ is weakly injective then this inclusion must split. So it suffices to prove (i) $\Rightarrow$ (ii), for which we show first that toposes of the form $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ are complete injectives and then that the class of complete injectives is closed under retracts.

For the first assertion, we recall from B3.2.5 that, for a cartesian internal category $\mathbb{C}$, $\mathbb{C}^{\mathrm{op}}$-torsors are the same thing as cartesian functors defined on $\mathbb{C}$; thus, for any geometric morphism $g \colon \mathcal{G} \to \mathcal{F}$ over $\mathcal{S}$, geometric morphisms $\mathcal{F} \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ over $\mathcal{S}$ correspond to cartesian $\mathcal{S}$-indexed functors $\mathbb{C} \to \mathbb{F}$, and the operation of composing a geometric morphism with g corresponds to that of composing a cartesian functor with (the indexed version of) $g^* \colon \mathbb{F} \to \mathbb{G}$. But the latter has an $\mathcal{S}$-indexed right adjoint, namely $g_* \colon \mathbb{G} \to \mathbb{F}$; so composition with the latter yields the functor g^+ which we seek. And the fact that, when g is an inclusion, we have $g^+(f)g \cong f$ is immediate from the definition of an inclusion.

Now suppose $\mathcal{E}$ is a retract of a complete injective $\tilde{\mathcal{E}}$. Given a geometric morphism $g \colon \mathcal{G} \to \mathcal{F}$, we may apply B1.1.10 to the diagram

$$\begin{tikzcd}
\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E}) \arrow[r] \arrow[d] & \mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\tilde{\mathcal{E}}) \arrow[r] \arrow[d] & \mathfrak{BTop}/\mathcal{S}\,(\mathcal{F},\mathcal{E}) \arrow[d] \\
\mathfrak{BTop}/\mathcal{S}\,(\mathcal{G},\mathcal{E}) \arrow[r] & \mathfrak{BTop}/\mathcal{S}\,(\mathcal{G},\tilde{\mathcal{E}}) \arrow[r] & \mathfrak{BTop}/\mathcal{S}\,(\mathcal{G},\mathcal{E})
\end{tikzcd}$$

where the vertical arrows are induced by composition with g and the horizontal ones by composition with the retraction morphisms $\mathcal{E} \leftrightarrows \tilde{\mathcal{E}}$. (As in the proof of 4.2.8, we have to regard this diagram as lying in the meta-2-category of Cauchy-complete categories; but this is permissible, by A4.1.15.) This yields the right adjoint g^+ for the topos $\mathcal{E}$, and the fact that the adjunction is a reflection if g is an inclusion. □

We pause to note a couple of elementary consequences of this result in the case $\mathcal{S} = \mathbf{Set}$:

Corollary 4.3.2 *Suppose $\mathcal{E}$ is an injective Grothendieck topos. Then*

(i) *$\mathcal{E}$ satisfies De Morgan's law (cf.* D4.6.2*);*

(ii) *the terminal object of $\mathcal{E}$ is an indecomposable projective (cf.* A1.1.10*).*

Proof (i) If $\mathcal{C}$ is cartesian, then the functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ satisfies De Morgan's law by D4.6.3(a); and De Morgan's law is inherited by retracts in $\mathfrak{Top}$, by D4.6.7(ii).

(ii) Again, the fact that the terminal object of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is an indecomposable projective when $\mathcal{C}$ is cartesian (or, more generally, when it has a terminal object) follows from A1.1.10; and it is easy to see that these properties are inherited by retracts in $\mathfrak{Top}$. □

Note that part (ii) of the corollary ensures that we have no hope of strengthening De Morgan's law to Booleanness in (i): the only Boolean injective topos (over **Set**) is **Set** itself.

The connection between injectivity and continuous categories, as studied in the last section, is given by the observation that the category of points of an injective topos must be continuous. (Note that in this section, by 'continuous category' we invariably mean 'continuous $\mathcal{S}$-indexed category'; but, in dealing with continuous categories, we shall continue to argue as if our base topos were **Set**.)

Given a bounded $\mathcal{S}$-topos $\mathcal{E}$, we shall write $\mathrm{Pt}(\mathcal{E})$ for the category $\mathfrak{BTop}/\mathcal{S}\,(\mathcal{S}, \mathcal{E})$ and call it the *category of points of* $\mathcal{E}$. By B3.4.8, we know that this category has filtered $\mathcal{S}$-indexed colimits. Moreover, the proof of B3.4.8 shows that, given a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between bounded $\mathcal{S}$-toposes, the functor $\mathrm{Pt}(\mathcal{F}) \to \mathrm{Pt}(\mathcal{E})$ induced by composition with f preserves filtered $\mathcal{S}$-indexed colimits. That is,

Lemma 4.3.3 *The assignment $\mathcal{E} \mapsto \mathrm{Pt}(\mathcal{E})$ defines a 2-functor from $\mathfrak{BTop}/\mathcal{S}$ to the meta-2-category $\mathfrak{FILT}_{\mathcal{S}}$ of $\mathcal{S}$-indexed categories with filtered colimits. functors preserving filtered colimits and (arbitrary) natural transformations.* □

Now if $\mathbb{C}$ is an internal category, we saw in B3.2.7 that the category $\mathrm{Pt}([\mathbb{C}, \mathcal{S}])$ is equivalent to the category $\mathbf{Tors}(\mathbb{C}) \subseteq [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ of $\mathbb{C}$-torsors; but, as we observed at the start of the last section, the latter is in turn equivalent to Ind-$\mathbb{C}$. We may thus conclude

Proposition 4.3.4 *Let $\mathcal{E}$ be a topos which is expressible as a retract in $\mathfrak{BTop}/\mathcal{S}$ of a functor category $[\mathbb{C}, \mathcal{S}]$. Then $\mathrm{Pt}(\mathcal{E})$ is an ind-small continuous category over $\mathcal{S}$. If $\mathcal{E}$ is injective in the sense of* 4.3.1, *then $\mathrm{Pt}(\mathcal{E})$ is also $\mathcal{S}$-cocomplete.*

Proof In the case when $\mathcal{E}$ is of the form $[\mathbb{C}, \mathcal{S}]$, then $\mathrm{Pt}(\mathcal{E}) \simeq$ Ind-$\mathbb{C}$ is continuous by 4.2.8, and ind-small by 4.2.18. The general case follows from 4.3.3 and 4.2.18. And if $\mathcal{E}$ is injective, then it can be expressed as a retract of $[\mathbb{C}, \mathcal{S}]$ where $\mathbb{C}$ is cocartesian, so that Ind-$\mathbb{C}$ and its retract $\mathrm{Pt}(\mathcal{E})$ are cocomplete by 4.2.21. □

But there is also a converse to 4.3.4. Recall from 4.2.2(iii) that, provided $\mathbb{C}$ is Cauchy-complete, it may be recovered up to equivalence from Ind-$\mathbb{C}$ as the full subcategory of finitely presentable objects. Hence the functor category $[\mathbb{C}, \mathcal{S}]$

(which, as we saw in A1.1.9, depends only on the Cauchy-completion of $\mathbb{C}$ – as does Ind-$\mathbb{C}$) may also be recovered from Ind-$\mathbb{C}$. In fact we have

Lemma 4.3.5 *For any internal category $\mathbb{C}$ in $\mathcal{S}$, $[\mathbb{C}, \mathcal{S}]$ is equivalent to the category $\mathfrak{FILT}_{\mathcal{S}}\,(\text{Ind-}\mathbb{C}, \mathbb{S})$.*

Proof This is just a particular case of the fact, established in 4.2.6, that Ind-$\mathbb{C}$ is the free filtered-colimit-completion of $\mathbb{C}$, since $\mathbb{S}$ has filtered $\mathcal{S}$-indexed colimits. □

We note that finite limits and arbitrary colimits in the category $\mathfrak{FILT}_{\mathcal{S}}\,(\text{Ind-}\mathbb{C}, \mathbb{S})$ may be constructed pointwise, from which it follows that, for any filtered-colimit-preserving functor $F\colon \text{Ind-}\mathbb{C} \to \text{Ind-}\mathbb{D}$, the functor $\mathfrak{FILT}_{\mathcal{S}}\,(\text{Ind-}\mathbb{D}, \mathbb{S}) \to \mathfrak{FILT}_{\mathcal{S}}\,(\text{Ind-}\mathbb{C}, \mathbb{S})$ induced by composition with F is geometric, and hence an inverse image functor by 2.2.10. That is, $\mathfrak{FILT}_{\mathcal{S}}\,(-, \mathbb{S})$ may be regarded as a functor $\mathfrak{K} \to \mathfrak{BTop}/\mathcal{S}$, where $\mathfrak{K}$ is the full sub-2-category of $\mathfrak{FILT}_{\mathcal{S}}$ whose objects are finitely accessible $\mathcal{S}$-indexed categories (i.e. categories of the form Ind-$\mathbb{C}$ with $\mathbb{C}$ small). On splitting the idempotents in $\mathfrak{K}$, we immediately deduce

Corollary 4.3.6 *Let $\mathbb{P}$ be an ind-small continuous $\mathcal{S}$-indexed category. Then there is a bounded $\mathcal{S}$-topos $\mathcal{E}$, determined up to equivalence by $\mathbb{P}$, which is a retract of a functor category $[\mathbb{C}, \mathcal{S}]$ and satisfies $\mathrm{Pt}(\mathcal{E}) \simeq \mathbb{P}$. If in addition $\mathbb{P}$ is $\mathcal{S}$-cocomplete, then $\mathcal{E}$ is injective.*

Proof We have seen that the 2-functors Pt and $\mathfrak{FILT}_{\mathcal{S}}\,(-, \mathbb{S})$ define an equivalence between the 2-category $\mathfrak{K}$ and the full sub-2-category of $\mathfrak{BTop}/\mathcal{S}$ whose objects are functor categories $[\mathbb{C}, \mathcal{S}]$; so they must extend to an equivalence between the Cauchy-completions of these two 2-categories. For the final assertion, we use 4.2.21 and 4.3.1(i). □

We note that, since every ind-small continuous category appears as a coadjoint retract of an object of $\mathfrak{K}$, it follows that every topos which occurs as a retract of a functor category $[\mathbb{C}, \mathcal{S}]$ is actually a coadjoint retract of such a category. This is clearly related to the equivalence of the injectivity notions considered in 4.3.1 – though it does not seem to be directly deducible from the latter, even in the case when $\mathbb{C}$ has finite colimits.

Suppose now that $\mathcal{E}$ is a retract in $\mathfrak{BTop}/\mathcal{S}$ of a functor category $[\mathbb{C}, \mathcal{S}]$. Then the retraction $r\colon [\mathbb{C}, \mathcal{S}] \to \mathcal{E}$ and the inclusion $i\colon \mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ compose to yield an idempotent endomorphism of $[\mathbb{C}, \mathcal{S}]$. However, the image of this endomorphism, in the sense of the factorization of A4.2.10, need not be equivalent to $\mathcal{E}$; nor need it be injective if $\mathbb{C}$ is cocartesian.

Example 4.3.7 Let $\mathcal{C}$ be a small category, A an object of $\mathcal{C}$ with more than one endomorphism, and let $E\colon \mathcal{C} \to \mathcal{C}$ be the constant functor with value A. Then E induces an endomorphism e of the topos $[\mathcal{C}, \mathbf{Set}]$, which is clearly idempotent since E is; but its image in the sense of A4.2.10 is the topos $[M, \mathbf{Set}]$ where

M is the monoid of all endomorphisms of A in $\mathcal{C}$, by A4.2.12(b) – and $[M, \mathbf{Set}]$ is not injective for any nontrivial M, since its terminal object is not projective (cf. 4.3.2(ii)). On the other hand, the image of e in the idempotent-splitting sense is simply **Set**, since the image of E in this sense is the terminal category **1**.

Since $\mathfrak{BTop}/\mathcal{S}$ has finite limits, we know that any idempotent 1-cell in this 2-category does have a splitting. Moreover, we have

Lemma 4.3.8 *Let $\mathcal{E}$ be a bounded $\mathcal{S}$-topos, and $e: \mathcal{E} \to \mathcal{E}$ a geometric morphism which carries an idempotent monad or comonad structure in $\mathfrak{BTop}/\mathcal{S}$ (cf.* B1.1.9(c)*). Then the image of e in the idempotent-splitting sense coincides with its image in the sense of* A4.2.10.

Proof We shall deal with the comonad case; the other is very similar. Let

$$\mathcal{E} \xrightarrow{\ r\ } \mathcal{F} \xrightarrow{\ i\ } \mathcal{E}$$

be a splitting of the idempotent e. Then, as we saw in B1.1.9(c), r is right adjoint to i in $\mathfrak{BTop}/\mathcal{S}$ (that is, r^* is left adjoint to i^*), and the unit $1_{\mathcal{F}} \to i^*r^*$ is an isomorphism. It follows that the counit $i^*i_* \to 1_{\mathcal{F}}$ of $(i^* \dashv i_*)$ is also an isomorphism, i.e. i is an inclusion. So the displayed diagram is also the surjection–inclusion factorization of e. □

Combining 4.3.4, 4.3.6 and 4.3.8, we have:

Theorem 4.3.9 *For a bounded $\mathcal{S}$-topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is a retract of an $\mathcal{S}$-topos of the form $[\mathbb{C}, \mathcal{S}]$.*

(ii) *$\mathcal{E}$ is a coadjoint retract of a topos of the form $[\mathbb{C}, \mathcal{S}]$.*

(iii) *There exists a local morphism $[\mathbb{C}, \mathcal{S}] \to \mathcal{E}$ over $\mathcal{S}$, for some internal category $\mathbb{C}$ in $\mathcal{S}$.*

Moreover, $\mathcal{E}$ satisfies the injectivity conditions of 4.3.1 *iff it satisfies any of the above conditions with the additional hypothesis that $\mathbb{C}$ is cocartesian.*

Proof The equivalence of (i) and (ii) follows as already indicated from 4.3.4, 4.3.6 and 4.2.18. Since the left adjoint of a local morphism is necessarily an inclusion by 3.6.1(iv), condition (iii) is equivalent to saying that $\mathcal{E}$ is the image in the sense of A4.2.10 of an idempotent comonad on $[\mathbb{C}, \mathcal{S}]$. So 4.3.8 implies that this is equivalent to (ii). The final assertion is immediate from 4.3.1. □

An $\mathcal{S}$-topos which is a retract of one of the form $[\mathbb{C}, \mathcal{S}]$ is called *quasi-injective*. Using 4.3.9, we may deduce various properties of such toposes: for example,

Corollary 4.3.10 *Let $p: \mathcal{E} \to \mathcal{S}$ be an $\mathcal{S}$-topos.*

(i) *If $\mathcal{E}$ is quasi-injective, then p is locally connected.*

(ii) *If $\mathcal{E}$ is injective, then p is totally connected.*

Proof (i) We already know that toposes of the form $[\mathbb{C}, \mathcal{S}]$ are locally connected over $\mathcal{S}$; so the result follows by applying 3.3.2(iv) to a triangle

$$\begin{tikzcd}
{[\mathbb{C}, \mathcal{S}]} \arrow[r, "f"] \arrow[dr] & \mathcal{E} \arrow[d, "p"] \\
& \mathcal{S}
\end{tikzcd}$$

where f is local (and hence connected).

(ii) In this case, we may choose $\mathbb{C}$ to be cocartesian and hence filtered; so the left adjoint of p^*, which is the composite $\varinjlim_{\mathbb{C}} f^*$ by the proof of 3.3.2(iv), is cartesian by B2.6.8. □

We remark that injective $\mathcal{S}$-toposes are also local over $\mathcal{S}$, by 3.6.3(b) and 3.6.4(ii). Indeed, they are actually hyperlocal (cf. B4.4.20(a)), and even carry algebra structures for the 'measure topos' monad of B4.5.8, as we observed in B4.5.10.

Remark 4.3.11 In connection with 4.3.9(iii), we note that $\mathcal{E}$ admits a totally connected morphism from $[\mathbb{C}, \mathcal{S}]$ iff it is an adjoint retract of $[\mathbb{C}, \mathcal{S}]$, by 4.3.8. However, not every quasi-injective $\mathcal{S}$-topos can be represented in this way. In fact, if there exists a totally connected morphism (or even an $\mathcal{S}$-essential surjection) $f\colon [\mathbb{C}, \mathcal{S}] \to \mathcal{E}$ over $\mathcal{S}$, then $\mathcal{E}$ must itself be of the form $[\mathbb{D}, \mathcal{S}]$, by 2.2.20: for $f_!$ preserves $\mathcal{S}$-indecomposable projectives since it has an $\mathcal{S}$-indexed right adjoint which preserves epimorphisms and $\mathcal{S}$-indexed coproducts, and the images under $f_!$ of the representable functors form a separating family for $\mathcal{E}$ since the right adjoint of $f_!$ is faithful (cf. B2.4.2(d)). Thus we have shown that all idempotent monads (unlike the idempotent comonads) in the 2-category $\mathfrak{D}\mathfrak{T}\mathsf{op}/\mathcal{S}$ of diagram toposes over $\mathcal{S}$ are already split.

To conclude this section, we give a characterization of quasi-injective toposes in terms of sites, which we shall require in the next section. Let $e\colon [\mathbb{C}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ be an idempotent geometric morphism. By the description on geometric morphisms between functor categories in terms of profunctors given in B3.2.8(c), we have an idempotent profunctor $\mathbb{W}\colon \mathbb{C} \nrightarrow \mathbb{C}$ such that $\mathbb{W}$ is a torsor for the right action of $\mathbb{C}$, which induces the functor f^* in the sense that for any discrete opfibration $\mathbb{F} \to \mathbb{C}$ we have $f^*(\mathbb{F}) \cong \mathbb{W} \otimes_{\mathbb{C}} \mathbb{F}$. If we think of the elements of $\mathbb{W}(U, V)$ as wavy arrows from U to V, as before, the condition that $\mathbb{W}$ should be a torsor is equivalent to saying that, for each V, the 'wavy slice category' $(\mathbb{C} \wr V)$ whose objects are the wavy arrows with codomain V, and whose morphisms are straight arrows between their domains forming commutative triangles, is filtered.

We now assume (as we are entitled to do, by 4.3.9) that $\mathbb{W}$ is actually an idempotent profunctor comonad. This means that each wavy arrow $f\colon U \rightsquigarrow V$

has an underlying straight arrow $\epsilon(f): U \to V$, and that ϵ preserves composition in the appropriate sense. Given such a $\mathbb{W}$, we define a coverage $J_{\mathbb{W}}$ on $\mathbb{C}^{\mathrm{op}}$ as follows: a cosieve on U is $J_{\mathbb{W}}$-covering iff it contains the underlying straight arrows of all wavy arrows with domain U. (Thus each object U has a smallest J-covering cosieve R_U, consisting precisely of the underlying straight arrows of all wavy arrows with domain U; however, $J_{\mathbb{W}}$ is not in general a rigid coverage in the sense of 2.2.18 – cf. 4.3.13 below.) It is clear that this definition satisfies the stability property (C) of 2.1.1; but it also satisfies the properties (M) and (L) of 2.1.8 (the latter since the idempotency of $\mathbb{W}$ means that every wavy arrow can be factored as the composite of two wavy arrows), so it is a Grothendieck coverage.

Lemma 4.3.12 *Under the above assumptions, the topos* $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$ *is the image (in either of the senses of* 4.3.8*) of the idempotent geometric morphism* $e: [\mathbb{C}, \mathcal{S}] \to [\mathbb{C}, \mathcal{S}]$ *corresponding to* $\mathbb{W}$.

Proof To identify the image of e in the sense of A4.2.10, we have to determine which subobjects $R \rightarrowtail \mathbb{C}(U, -)$ (i.e. which cosieves on an object U of $\mathbb{C}$) are mapped to isomorphisms by the functor $\mathbb{W} \otimes_{\mathbb{C}} (-)$. But

$$\mathbb{W} \otimes_{\mathbb{C}} R \rightarrowtail \mathbb{W} \otimes_{\mathbb{C}} \mathbb{C}(U, -) \cong \mathbb{W}(U, -)$$

is an isomorphism iff, for each wavy arrow $f: U \rightsquigarrow V$, there exists $g: U \to W$ in R and a wavy arrow $h: W \rightsquigarrow V$ such that $h \cdot g = f$. Clearly, this condition implies that R contains $\epsilon(f) = \epsilon(h) \cdot g$ for any such f; but conversely, if R contains all the $\epsilon(f)$, then we may factor any wavy $f: U \rightsquigarrow V$ as a composite hg of two wavy arrows, and then $\epsilon(g) \in R$, so the condition is satisfied. Thus the image of e is precisely $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$.

By 4.3.8, we know that $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$ is also the image of e in the idempotent-splitting sense. However, we may alternatively prove this by a direct calculation, which we shall in any case need in the next section (cf. the proof of 4.4.4). It is sufficient to show that if $F: \mathbb{C} \to \mathbb{S}$ is any $J_{\mathbb{W}}$-sheaf then $e_*(F) \cong F$. But since each object of $\mathbb{C}$ has a smallest $J_{\mathbb{W}}$-covering sieve, the assertion that F is a $J_{\mathbb{W}}$-sheaf means simply that, for each $U \in C_0$, the canonical morphism $F(U) \to \varprojlim_{f: U \to V \in R_U} F(V)$ is an isomorphism. And $e_*(F)$ is by definition $\mathbb{W} \pitchfork_{\mathbb{C}} F$, where $\mathbb{W} \pitchfork_{\mathbb{C}} (-)$ denotes the right adjoint of $\mathbb{W} \otimes_{\mathbb{C}} (-)$; from this, we may readily deduce the formula $e_*(F)(U) = \varprojlim F(V)$, where this limit is taken (not over R_U but) over the category $(U \wr \mathbb{C})$ whose objects are wavy arrows $U \rightsquigarrow V$ and whose morphisms are straight arrows between their codomains inducing commutative triangles. It is clear that the canonical mapping from the first limit above to the second (which sends $(x_f \mid f \in R_U)$ to $(x_{\epsilon(f)} \mid f: U \rightsquigarrow V)$) is monic; to show that it is an isomorphism, we have to show that if $(x_f \mid f: U \rightsquigarrow V)$ is any element of the second limit, and f and g are two wavy arrows with the same underlying straight arrow, then necessarily $x_f = x_g$. But if f and g have the same underlying straight arrow, then they are coequalized by any wavy arrow

whose domain matches their codomain (V, say), and hence x_f and x_g have the same image in $e_*(F)(V)$. Since we have already observed that $F \to e_*(F)$ is monic, this forces $x_f = x_g$, as required. □

It follows from 4.3.12 that the idempotent endomorphism e of $[\mathbb{C}, \mathcal{S}]$, and hence the idempotent profunctor comonad $\mathbb{W}$, depend only on the image of the counit map ϵ, since the knowledge of which morphisms of $\mathbb{C}$ underlie wavy arrows suffices to define the coverage $J_{\mathbb{W}}$. However, it does not seem easy to give a direct construction of $\mathbb{W}$ from $J_{\mathbb{W}}$, except in the particular case when ϵ is monic (in which case, of course, $\mathbb{W}(U, V)$ may be taken to be the set of those morphisms $U \to V$ which belong to R_U).

Example 4.3.13 If P is a continuous poset (in **Set**, for simplicity), then 4.3.12 allows us to express it as the category of points of a coadjoint retract of the topos $[P, \mathbf{Set}]$: specifically, of $\mathbf{Sh}(P^{\mathrm{op}}, J_{<<})$, where $J_{<<}$ is the (Grothendieck) coverage in which a cosieve on an element p of P is covering iff it contains all elements q with $p << q$. For example, if P is the ordered set $(-\infty, 0] \subseteq \mathbb{R}$ (which is easily seen to be continuous, since its way-below relation coincides with the strict ordering $<$), we obtain precisely the non-rigid coverage on $[0, \infty) \cong (-\infty, 0]^{\mathrm{op}}$ mentioned as a counterexample after 2.2.18. On the other hand, if P is the poset IQ of ideals of a poset Q, then we obtain the rigid coverage associated (as in 2.2.4(d)) with the inclusion of the sub-poset of principal ideals.

Suggestions for further reading: Johnstone [517], Johnstone & Joyal [542], Scott [1102].

C4.4 Exponentiable toposes

We now turn from the topos analogue of 4.1.6 to that of 4.1.9: the determination of which bounded $\mathcal{S}$-toposes are exponentiable in the 2-category $\mathfrak{BTop}/\mathcal{S}$. We recall that we made some preliminary investigations of exponentiability in Section B4.3; in particular, we recall the following result, which we restate as the starting-point of our present investigation:

Theorem 4.4.1 *Let $p: \mathcal{E} \to \mathcal{S}$ be a bounded $\mathcal{S}$-topos. Then $\mathcal{E}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$ iff the particular exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$ exists, where $\mathcal{S}[\mathbb{O}]$ is the object classifier for toposes defined over $\mathcal{S}$ (cf.* D3.2.1*).* □

This of course corresponds to the fact, which we observed in the proof of 4.1.9, that the existence of the particular exponential S^X, where S is the Sierpiński space, implies the exponentiability of X in **Loc**. And we shall use it in much the same way, to establish the link between injectivity and exponentiability in $\mathfrak{BTop}/\mathcal{S}$. Note that $\mathcal{S}[\mathbb{O}]$ is an injective $\mathcal{S}$-topos by 4.3.1, since we may identify it with the functor category $[\mathbb{S}_f, \mathcal{S}]$ and $\mathbb{S}_f$ is cocartesian; and since the functor $(-) \times \mathcal{E}: \mathfrak{BTop}/\mathcal{S} \to \mathfrak{BTop}/\mathcal{S}$ preserves inclusions by the proof of B3.3.6, it

follows that its right adjoint, if it exists, must preserve injectives. But we have

$$\mathrm{Pt}(\mathcal{S}[\mathbb{O}]^{\mathcal{E}}) \simeq \mathfrak{BTop}/\mathcal{S}\,(\mathcal{E}, \mathcal{S}[\mathbb{O}]) \simeq \mathcal{E}$$

by the defining property of the object classifier; so from 4.3.4 we immediately deduce

Proposition 4.4.2 *If $\mathcal{E}$ is an exponentiable $\mathcal{S}$-topos, then it is continuous as an $\mathcal{S}$-indexed category.* □

Our main task in this section is to establish the converse of 4.4.2. We know from 2.2.8 that any bounded $\mathcal{S}$-topos $\mathcal{E}$ has a small dense subcategory, which may be taken to be closed under finite colimits in $\mathcal{E}$; so if $\mathcal{E}$ is a continuous category then it is necessarily ind-small, as well as cocomplete. Hence by 4.3.6 it is (equivalent to) the category of points of an injective $\mathcal{S}$-topos $\mathcal{F}$, which is determined up to equivalence by $\mathcal{E}$, and which is clearly the only possible candidate for the exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$. As in the proof of 4.1.9, the hard work consists in showing that it does indeed have the universal property of this exponential.

Before we embark on it, let us fix some notation. We shall take $\mathbb{C}$ to be a small full generating subcategory of $\mathbb{E}$, which is closed under finite limits as well as colimits, and is therefore a pretopos. We shall write K for the Grothendieck coverage on $\mathbb{C}$ induced by the inclusion $\mathbb{C} \to \mathbb{E}$, i.e. the collection of all sieves on $\mathbb{C}$ which map to epimorphic families in $\mathbb{E}$. We shall write $\mathbb{W}$ both for the 'wavy arrow' profunctor $\mathcal{E} \nrightarrow \mathcal{E}$ and for its restriction to $\mathbb{C}$, and $J_{\mathbb{W}}$ for the coverage which it induces on $\mathbb{C}^{\mathrm{op}}$ as in the last section. (Thus $\mathcal{F}$ is the topos $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$.)

We also need a description of K in terms of $\mathbb{W}$. Note that, as usual, K contains the coherent coverage P on $\mathbb{C}$, and so by 3.2.18 it may be expressed as the join $P \vee K_d$, where K_d consists of all the dm-sieves in K. But from the definition of wavy arrow, every K_d-covering sieve on an object U of $\mathbb{C}$ contains the underlying straight arrows of all wavy arrows with codomain U. The sieve S_U consisting of all such straight arrows is not in general a dm-sieve, or even filtered (though it will be filtered if the 'underlying straight arrow' map is injective), but it is K-covering since $\mathcal{E}$ is a continuous category. So we see that K may be generated as a Grothendieck coverage by P together with the coverage $K_{\mathbb{W}}$ whose only covers are the sieves S_U, $U \in C_0$. (Note that the latter is indeed a coverage, since if $f\colon V \to U$ in $\mathbb{C}$ then $S_V \subseteq f^*S_U$. Indeed, if we modify the definition to include all sieves which contain S_U for some U, we get a Grothendieck coverage; the 'local character' axiom (L) follows from the idempotency of $\mathbb{W}$, as it did for $J_{\mathbb{W}}$.) In particular, a functor $\mathbb{C}^{\mathrm{op}} \to \mathbb{S}$ is a K-sheaf (that is, an object of $\mathcal{E}$) iff it is a P-sheaf and satisfies the sheaf axiom for the sieves S_U.

Our basic strategy for proving the exponentiability of $\mathcal{E}$ is to show that the equivalence between $\mathcal{E}$ (that is, $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, K)$) and the category of points of $\mathcal{F}$ (i.e. of cartesian $J_{\mathbb{W}}$-cover-preserving functors $\mathbb{C}^{\mathrm{op}} \to \mathbb{S}$) can be derived from properties of the pair $(\mathbb{C}, \mathbb{W})$ which are preserved by inverse image functors.

Then for any $\mathcal{S}$-topos $r\colon \mathcal{G} \to \mathcal{S}$ we shall have

$$\begin{aligned}\mathfrak{BTop}/\mathcal{S}\,(\mathcal{G} \times_{\mathcal{S}} \mathcal{E}, \mathcal{S}[\mathbb{O}]) &\simeq \mathcal{G} \times_{\mathcal{S}} \mathcal{E} \\ &\simeq \mathbf{Sh}_{\mathcal{G}}(r^*\mathbb{C}, r^\# K) \\ &\simeq \mathfrak{BTop}/\mathcal{G}\,(\mathcal{G}, \mathbf{Sh}_{\mathcal{G}}(r^*\mathbb{C}^{\mathrm{op}}, r^\# J_{\mathbb{W}})) \\ &\simeq \mathfrak{BTop}/\mathcal{S}\,(\mathcal{G}, \mathcal{F})\end{aligned}$$

yielding the required adjunction. (It should be noted, however, that $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, K)$ and the category of cartesian $J_{\mathbb{W}}$-cover-preserving functors $\mathbb{C}^{\mathrm{op}} \to \mathcal{S}$ do not coincide as subcategories of $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$; indeed, as subcategories of Ind-$\mathbb{C} \simeq \mathfrak{Cart}_{\mathcal{S}}\,(\mathbb{C}^{\mathrm{op}}, \mathcal{S})$, they are embedded via the right and left adjoints respectively to the colimit functor Ind-$\mathbb{C} \to \mathbb{E}$.)

One further ingredient is needed before we can actually prove such a theorem. In general, if $\mathbb{C}$ is a pretopos, then functors on $\mathbb{C}^{\mathrm{op}}$ which are sheaves for the coherent coverage are 'almost cartesian' in the sense of the remark before A1.4.8, but they need not be cartesian: they send coequalizers of equivalence relations to equalizers, but they do not necessarily send arbitrary coequalizers to equalizers, since they do not preserve the infinite unions involved in the definition of the equivalence relation generated by a parallel pair. Since we wish this to be true in our case, we have to build it into the properties of $\mathbb{W}$, as follows. We shall say that the pair $(\mathbb{C}, \mathbb{W})$ is *essentially finitary* if, given any object U of $\mathbb{C}$ and any relation $T \rightarrowtail U \times U$, with equivalence closure $R = \bigcup_{n=1}^{\infty} T^{(n)}$ defined as in 3.4.5, each wavy arrow $V \rightsquigarrow R$ in $\mathbb{C}$ factors through $T^{(n)} \rightarrowtail R$ for some finite n. Clearly, if $(\mathbb{C}, \mathbb{W})$ is obtained as above from a topos $\mathcal{E}$ which is a continuous category, then it is essentially finitary; for the $T^{(n)}$ then form a directed diagram in $\mathcal{E}$ whose colimit is R, and so there must be a morphism of ind-objects from $w(R)$ to this diagram.

The key 'change of base' result is thus

Lemma 4.4.3 *Let $\mathbb{C}$ be an internal pretopos (with arbitrary coequalizers) in a topos $\mathcal{S}$, and $\mathbb{W}\colon \mathbb{C} \nrightarrow \mathbb{C}$ an idempotent profunctor comonad which is a right torsor, such that the pair $(\mathbb{C}, \mathbb{W})$ is essentially finitary. Then, for any geometric morphism $q\colon \mathcal{F} \to \mathcal{S}$, the pair $(q^*\mathbb{C}, q^*\mathbb{W})$ is essentially finitary.*

Proof The assertion that $(\mathbb{C}, \mathbb{W})$ is essentially finitary may be expressed as a geometric sequent in the internal language of $\mathcal{S}$, provided the infinite union $\bigcup_{n=1}^{\infty} T^{(n)}$ always exists in $\mathbb{C}$. But, since we have assumed that $\mathbb{C}$ has arbitrary coequalizers, we may construct it as the kernel-pair of the coequalizer of $T \rightrightarrows U$. These constructions are preserved by the inverse image functor q^*; hence it preserves the validity of the geometric sequent. □

Proposition 4.4.4 *Let $\mathbb{C}$ be an internal pretopos with arbitrary coequalizers in a topos $\mathcal{S}$, and let $\mathbb{W}$ be an idempotent profunctor comonad on $\mathbb{C}$ which is a right torsor, such that the pair $(\mathbb{C}, \mathbb{W})$ is essentially finitary. Let K be the coverage*

on $\mathbb{C}$ generated by the coherent coverage together with the sieves S_U, and let $J_{\mathbb{W}}$ be the coverage on $\mathbb{C}^{\mathrm{op}}$ consisting of the sieves R_U defined in 4.3.12. *Then* $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, K)$ *is equivalent to the category of points of the topos* $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$.

Proof The 'essentially finitary' assumption ensures that any K-sheaf on $\mathbb{C}$, considered as a functor $\mathbb{C}^{\mathrm{op}} \to \mathbb{S}$, preserves equalizers, since it implies that the kernel-pair of the coequalizer of an arbitrary pair $T \rightrightarrows U$ is K-covered by the sieve generated by the morphisms $T^{(n)} \rightarrowtail R$. So we may regard $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}, K)$ as a subcategory of the category $\mathfrak{Cart}(\mathbb{C}^{\mathrm{op}}, \mathcal{S}) \simeq$ Ind-$\mathbb{C}$ – in fact a reflective subcategory thereof, since it is reflective in $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$.

On the other hand, the category of points of $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$ may also be regarded as a full subcategory of $\mathfrak{Cart}(\mathbb{C}^{\mathrm{op}}, \mathcal{S})$, consisting of those cartesian functors which map $J_{\mathbb{W}}$-covering sieves to epimorphic families. As we have already remarked, these two subcategories of $\mathfrak{Cart}(\mathbb{C}^{\mathrm{op}}, \mathcal{S})$ do not coincide in general. However, since $\mathbb{W}$ is a right torsor, the functor $(-) \otimes_{\mathbb{C}} \mathbb{W} \colon [\mathbb{C}^{\mathrm{op}}, \mathcal{S}] \to [\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$ maps the subcategory $\mathfrak{Cart}(\mathbb{C}^{\mathrm{op}}, \mathcal{S})$ into itself; and it inherits an idempotent comonad structure from $\mathbb{W}$. By 4.3.12, we know that its image is exactly the subcategory of functors which preserve covers in $J_{\mathbb{W}}$, since it corresponds to the image of the idempotent endomorphism of $[\mathbb{C}, \mathcal{S}]$ induced by $\mathbb{W}$.

As a functor on $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$, $(-) \otimes_{\mathbb{C}} \mathbb{W}$ has a right adjoint $\mathbb{W} \pitchfork^{\mathbb{C}} (-)$, by B2.7.8. The latter may be defined more explicitly by

$$\mathbb{W} \pitchfork^{\mathbb{C}} F(U) = \varprojlim F(V),$$

where the limit is taken over the category $(\mathbb{C} \wr U)$ whose objects are all wavy arrows $V \rightsquigarrow U$ and whose morphisms are straight arrows between their domains. Clearly, $\mathbb{W} \pitchfork^{\mathbb{C}} (-)$ has an idempotent monad structure, and, by the argument in the proof of 4.3.12 (applied with all arrows reversed), its image (as a functor on $[\mathbb{C}^{\mathrm{op}}, \mathcal{S}]$) consists precisely of those functors $\mathbb{C}^{\mathrm{op}} \to \mathcal{S}$ which satisfy the sheaf axiom for the sieves S_U defined earlier. But the fact that $\mathbb{W}$ is a right torsor ensures that this functor also preserves sheaves for the coherent coverage, since the latter can be defined as functors which preserve certain finite limits; so its image as a functor on $\mathfrak{Cart}(\mathbb{C}^{\mathrm{op}}, \mathcal{S})$ is exactly the category of K-sheaves. Moreover, the adjunction between $(-) \otimes_{\mathbb{C}} \mathbb{W}$ and $\mathbb{W} \pitchfork^{\mathbb{C}} (-)$ is itself idempotent; so it restricts to an equivalence between the category of K-sheaves and the category of cartesian functors which preserve $J_{\mathbb{W}}$-covering sieves. □

Theorem 4.4.5 *A bounded $\mathcal{S}$-topos $p \colon \mathcal{E} \to \mathcal{S}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$ iff its underlying category is continuous (as an $\mathcal{S}$-indexed category).*

Proof One direction is 4.4.2. For the converse, suppose $\mathcal{E}$ is a continuous category; then it suffices by 4.4.1 to construct the exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$. We define the latter to be $\mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})$, where $\mathbb{C}$ is a generating subcategory for $\mathcal{E}$ closed under finite limits and colimits, $\mathbb{W}$ is the restriction to $\mathbb{C}$ of the wavy-arrow profunctor on $\mathcal{E}$, and $J_{\mathbb{W}}$ is the coverage defined from $\mathbb{W}$ as in 4.3.12. We note

that $\mathbb{W}$ is a right torsor, since $\mathbb{C}$ is closed under finite colimits in $\mathcal{E}$, and by 4.2.17 these are also 'wavy colimits', i.e. $\mathbb{W}(-,V)$ preserves finite limits for any V.

Now let $q: \mathcal{F} \to \mathcal{S}$ be an arbitrary $\mathcal{S}$-topos. Then the pair $(q^*\mathbb{C}, q^*\mathbb{W})$ is essentially finitary by 4.4.3; moreover, the constructions of the coverages $J_{\mathbb{W}}$ and $K_{\mathbb{W}}$ from $\mathbb{W}$ make it clear that we have $q^{\#}(J_{\mathbb{W}}) = J_{q^*\mathbb{W}}$ and $q^{\#}(K_{\mathbb{W}}) = K_{q^*\mathbb{W}}$. So, applying 4.4.4 in the context of $\mathcal{F}$-indexed categories, we have

$$\begin{aligned}\mathfrak{BTop}/\mathcal{S}\,(\mathcal{F} \times_{\mathcal{S}} \mathcal{E}, \mathcal{S}[\mathbb{O}]) &\simeq \mathcal{F} \times_{\mathcal{S}} \mathcal{E} \\ &\simeq \mathbf{Sh}_{\mathcal{F}}(q^*\mathbb{C}, q^{\#}K) \\ &\simeq \mathfrak{BTop}/\mathcal{F}\,(\mathcal{F}, \mathbf{Sh}_{\mathcal{F}}(q^*\mathbb{C}^{\mathrm{op}}, q^{\#}J_{\mathbb{W}})) \\ &\simeq \mathfrak{BTop}/\mathcal{S}\,(\mathcal{F}, \mathbf{Sh}_{\mathcal{S}}(\mathbb{C}^{\mathrm{op}}, J_{\mathbb{W}})),\end{aligned}$$

as required. □

Examples 4.4.6 (*a*) In Section B4.3, we were mainly concerned with two particular classes of exponentiable $\mathcal{S}$-toposes, namely those of the form $[\mathbb{C}, \mathcal{S}]$ for an (arbitrary) internal category $\mathbb{C}$ in $\mathcal{S}$ and those which are coherent over $\mathcal{S}$. It is easy to see that both classes consist of exponentiable toposes using 4.4.5, since the underlying categories of both types are locally finitely presentable (for the local finite presentability of coherent toposes, cf. D3.3.12).

(*b*) From 4.2.9, we may deduce that if $\mathcal{E}$ is an exponentiable $\mathcal{S}$-topos, then so is $\mathcal{E}/A$ for any A. (However, this could also be deduced from (*a*) and the composability of exponentiable morphisms, as we saw in B4.3.6.)

(*c*) As a special case of (*b*), any open subtopos of an exponentiable topos is exponentiable. But any closed subtopos $\mathcal{E}'$ of an exponentiable topos $\mathcal{E}$ is also exponentiable by 4.2.8, since the inclusion $\mathcal{E}' \to \mathcal{E}$ preserves filtered colimits by (the $\mathcal{S}$-indexed version of) A4.5.4, and so does its left adjoint. (Like (*b*), this could alternatively have been deduced from the second case of (*a*) and the composability of exponentiable morphisms; for a closed subtopos of $\mathcal{E}$ is coherent as an $\mathcal{E}$-topos.) Combining the two, we deduce that any locally closed subtopos (that is, any subtopos which is the intersection of an open and a closed subtopos) of an exponentiable topos is exponentiable. Conversely, if an inclusion $\mathcal{S}' \to \mathcal{S}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, then it is locally closed; this follows from (*f*) below and 4.1.12, since the terminal object of $\mathbf{Loc}(\mathcal{S})$ is compact and Hausdorff.

(*d*) Paralleling 4.1.7(ii), if a geometric morphism $f: \mathcal{F} \to \mathcal{E}$ over $\mathcal{S}$ is connected and either locally connected or tidy, and $\mathcal{F}$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, then $\mathcal{E}$ is also exponentiable, since we can use f^* and its left (resp. right) adjoint to express $\mathcal{E}$ as a retract of $\mathcal{F}$ by functors preserving filtered colimits. (However, exponentiability does not 'descend' along arbitrary open or proper surjections, as we shall see in 4.4.9 below.)

(*e*) We observed in B4.3.5(iii) that any retract in $\mathfrak{BTop}/\mathcal{S}$ of an exponentiable $\mathcal{S}$-topos is exponentiable; but we may alternatively derive this from 4.4.5 and 4.2.8, since inverse image functors preserve filtered colimits. In particular,

we note that every quasi-injective topos (in the sense of 4.3.9) is exponentiable. So, if we can construct the exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$, we can also construct $\mathcal{S}[\mathbb{O}]^{(\mathcal{S}[\mathbb{O}]^{\mathcal{E}})}$, and so on. In the terminology introduced in B4.3.2, this means that the 'dual' of any dualizable geometric theory is itself dualizable.

(*f*) If a localic topos $\mathbf{Sh}_{\mathcal{S}}(X)$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, then it follows from 4.4.2 and 4.2.11 that X must be a locally compact locale. (We could also have deduced this by noting that $\mathrm{Sub}_{\mathbf{Sh}(X)}(1)$ may be identified with the category of points of the exponential $\mathbf{Sh}(S)^{\mathbf{Sh}(X)}$, where S is the Sierpiński space, and the latter is an injective topos since $\mathbf{Sh}(S)$ is injective.) Somewhat surprisingly, in view of the close formal similarity between 4.1.9 and 4.4.5, this necessary condition for exponentiability of a localic topos is not sufficient; we shall see a counterexample in 4.4.8 below. However, if X is stably locally compact, then $\mathbf{Sh}_{\mathcal{S}}(X)$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$, by 4.1.13 and (*e*) above. In fact a stronger result is true: if X is locally stably compact, then $\mathbf{Sh}_{\mathcal{S}}(X)$ is exponentiable. (This follows from the remark after 4.1.13, plus the (easy) fact that any locally coherent topos is an open subtopos of a coherent topos: given a distributive pre-lattice B, let B^+ be the lattice obtained by adjoining a top element to B, and then IB is a principal ideal in $I(B^+)$.)

To determine exactly which locales give rise to exponentiable toposes, we need to introduce a strengthening of the way-below relation, which stands in much the same relation to it as the notion of strong compactness introduced in 3.4.1 does to ordinary compactness. Given elements U and V of a frame $\mathcal{O}(X)$, we shall write $V \prec\!\!\prec U$ if, whenever we have a filtered diagram $(A_i \mid i \in I)$ in $\mathbf{Sh}(X)$ with $\varinjlim_I A_i \cong U$, there exists $i \in I$ such that A_i has a section over V. And we shall say that X is *locally metastably compact* if every $U \in \mathcal{O}(X)$ satisfies $U = \bigcup\{V \in \mathcal{O}(X) \mid V \prec\!\!\prec U\}$. (The reason for the term 'metastably' is that, in a locally stably compact locale, the relation $\prec\!\!\prec$ coincides with $<<$, and hence any such locale is locally metastably compact. We shall not prove this here, since we have already seen that locally stably compact locales are exponentiable.)

If we already know that $\mathbf{Sh}(X)$ is a continuous category, then it clearly suffices in the definition of the relation $\prec\!\!\prec$ to consider the particular filtered diagram $w(U)$; hence in this case we have $V \prec\!\!\prec U$ iff there exists a wavy arrow $V \rightsquigarrow U$ in $\mathbf{Sh}(X)$. Thus we may conclude

Lemma 4.4.7 *If X is a locale such that $\mathbf{Sh}(X)$ is an exponentiable $\mathcal{S}$-topos, then it is locally metastably compact.*

Proof Given $U \in \mathcal{O}(X)$, we identify it with a subterminal object of $\mathbf{Sh}(X)$ and let $w(U) = (A_i \mid i \in I)$. Then the supports σA_i of the A_i form a directed family of subterminal objects whose union is U, and each σA_i can be covered by opens V over which A_i admits a section. Thus U is the union of all such V. □

Example 4.4.8 We may now give our promised example of a locally compact locale which is not locally metastably compact, and hence does not give rise to an exponentiable topos of sheaves. The space X of Example 4.1.10, though not locally stably compact, is locally metastably compact, since although the relation $X \prec\!\!\prec X$ fails to hold (as we saw in 3.4.1(*a*)), we do have $U \prec\!\!\prec X$ and $V \prec\!\!\prec X$, where U and V are the two copies of $[0,1]$ embedded in X. However, if we 'iterate' this construction as follows, we obtain our desired counterexample.

Consider the product space $[0,1] \times K$, where K is the Cantor space, i.e. the product of countably many copies of the discrete two-point space. (We think of points of K as infinite sequences of 0's and 1's.) We define an equivalence relation R on this space by setting $(x,s)\,R\,(y,t)$ iff $x = y$, and either $s = t$ or there exists a natural number n such that $x > 1/2^{n+1}$ and the first n terms of the sequences s and t agree; let Y be the quotient space $([0,1] \times K)/R$. Thus, in terms of its evident projection onto $[0,1]$, Y has a discrete 2^n-point space as its fibre over each point of the interval $(1/2^{n+1}, 1/2^n]$, and a copy of K as its fibre over 0. It is easy to see that the quotient map $q\colon [0,1] \times K \to Y$ is open, and hence by 4.1.7(ii) Y is locally compact. But we shall show that no open $U \subseteq Y$ satisfying $U \prec\!\!\prec Y$ can meet the fibre over 0; hence Y is not locally metastably compact.

Let U be an open set in Y containing the point $q(0,s)$. Then, for some n, U also contains the points $q(1/2^{n+1}, s)$ and $q(1/2^{n+1}, s')$ where s' differs from s at the $(n+1)$st term but not before. Hence U contains an open subset U' which looks like the effect of identifying two copies of the open interval $(t-\epsilon, t+\epsilon)$ along the subinterval $(t, t+\epsilon)$, where $t = 1/2^{n+1}$ and $0 < \epsilon < 1/2^{n+2}$. For $0 < \delta < \epsilon$, let X_δ be the space obtained from X by 'unglueing' these two intervals over $(t, t+\delta]$. Then the projection $X_\delta \to X$ is a local homeomorphism, and it is clear that we can regard $(X_\delta \mid \delta > 0)$ as a directed diagram of sheaves on X with colimit X. But no X_δ admits a section over U', let alone over U.

Remark 4.4.9 The fact that the space Y of 4.4.8 is an open surjective image of the compact Hausdorff (and therefore stably locally compact) space $[0,1] \times K$ shows us that *exponentiability does not descend along open surjections*, in contrast to B4.3.7: that is, if we are given an open surjection $\mathcal{F} \to \mathcal{E}$ in $\mathfrak{BTop}/\mathcal{S}$ and $\mathcal{F}$ is exponentiable, we cannot conclude that $\mathcal{E}$ is exponentiable. At first sight, this is rather surprising, since the descent theorem for open surjections, to be proved in the next chapter (see 5.1.6), tells us that in these circumstances $\mathcal{E}$ is the pullback-stable pseudo-colimit in $\mathfrak{BTop}/\mathcal{S}$ of a diagram involving $\mathcal{F}$ and the pullback $\mathcal{F} \times_{\mathcal{E}} \mathcal{F}$; and we might expect to be able to form the exponential $\mathcal{S}[\mathbb{O}]^{\mathcal{E}}$ as a corresponding pseudo-limit of exponentials. The reason why this will not work is that, in contrast to the special case considered in B4.3.7, the kernel-pair $\mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ does not in general inherit exponentiability from $\mathcal{F}$: in the particular example just described, the equivalence relation R is (not locally closed in $([0,1] \times K) \times ([0,1] \times K)$, and hence) not locally compact.

To prove the converse of 4.4.7, let X be a locally metastably compact locale. It suffices by 4.2.10 to construct an initial ind-object in $\mathbf{Sh}(X)$ with colimit E, for every sheaf E on X. But it is easy to see that local metastable compactness is a local property, i.e. it is inherited by the domain of any local homeomorphism $E \to X$; so in fact it suffices to do this for the particular case $E = X$. In what follows, we shall freely identify sheaves on X with local homeomorphisms over X.

We say that a sheaf $F \to X$ is *admissible* if it is generated by a finite number of sections (i.e. there exists an epimorphism $\coprod_{i=1}^{n} U_i \twoheadrightarrow F$ in $\mathbf{Sh}(X)$ for some n, where the U_i are subterminal objects of $\mathbf{Sh}(X)$) and, for every ind-object $(G_j \mid j \in J)$ in $\mathbf{Sh}(X)$ with colimit X, there exists a morphism $F \to G_j$ for some j. (Thus a subterminal object U is admissible iff it satisfies $U \prec\!\!\prec X$.) We define a category $\mathcal{I}$ as follows: its objects are monomorphisms $(F' \rightarrowtail F)$ where F is admissible and $F' \ll F$ in the (continuous) lattice of subsheaves of F, and morphisms $(F' \rightarrowtail F) \to (G' \rightarrowtail G)$ are morphisms $F' \to G'$ for which there exists some morphism $F \to G$ making the obvious square commute.

Lemma 4.4.10 *The category $\mathcal{I}$ is filtered.*

Proof It is clearly nonempty. Given two admissible objects F and G, it is easy to verify that their coproduct $F \amalg G$ is admissible; also, if $F' \ll F$ and $G' \ll G$, then we have $F' \ll F \leq F \amalg G$ and $G' \ll G \leq F \amalg G$ in $\mathrm{Sub}(F \amalg G)$, and hence $F' \amalg G' \ll F \amalg G$. So $(F' \amalg G' \rightarrowtail F \amalg G)$ is an object of $\mathcal{I}$, and it clearly admits morphisms from $(F' \rightarrowtail F)$ and $(G' \rightarrowtail G)$.

Now suppose given a parallel pair of morphisms $f', g' : (F' \rightarrowtail F) \rightrightarrows (G' \rightarrowtail G)$ in $\mathcal{I}$. Form the diagram

$$\begin{tikzcd}
F' \arrow[r, shift left, "f'"] \arrow[r, shift right, "g'"'] \arrow[d, tail] & G' \arrow[r, two heads, "h'"] \arrow[d, tail] & H' \arrow[d] \\
F \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & G \arrow[r, "h"] & H
\end{tikzcd}$$

in $\mathbf{Sh}(X)$, where the top row is a coequalizer and the right-hand square is a pushout. By A2.4.3, the morphism $H' \to H$ is monic, and the square is also a pullback; and since h is epic (so that pullback along it is a conservative functor) it is easy to verify that we have $H' \ll H$ in $\mathrm{Sub}(H)$. It is also clear that H is finitely generated; to show that it is admissible, consider an ind-object $(A_j \mid j \in J)$ with colimit X. For some i, we have a morphism $k : G \to A_i$. The composites kf and kg need not be equal; but since $F' \ll F$ in $\mathrm{Sub}(F)$ we can find $j \to j'$ in J such that the equalizer of $F \rightrightarrows G \to A_j \to A_{j'}$ contains $F' \rightarrowtail F$. and hence the composite $G' \rightarrowtail G \to A_j \to A_{j'}$ factors through the coequalizer h'. So we obtain a factorization of $G \to A_j \to A_{j'}$ through the pushout h; hence

H is admissible, and $h' : (G' \rightarrowtail G) \to (H' \rightarrowtail H)$ is a morphism of $\mathcal{I}$ coequalizing f' and g'. □

The finite generation condition in the definition of admissible objects ensures that $\mathcal{I}$ is at least equivalent to a small category. Thus we may regard the obvious forgetful functor $\mathcal{I} \to \mathbf{Sh}(X)$ sending $(F' \rightarrowtail F)$ to F' as an ind-object in $\mathbf{Sh}(X)$.

Lemma 4.4.11 *The ind-object just defined is initial among ind-objects in* $\mathbf{Sh}(X)$ *with colimit* X.

Proof First we must verify that the ind-object does indeed have colimit X. But since X can be covered by opens U' for which there exists U with $U' << U \prec\!\prec X$, it is clear that the colimit is well-supported. To show that it is subterminal, suppose $U \prec\!\prec X$ and we are given morphisms $f, g : U \rightrightarrows F'$ for some object $(F' \rightarrowtail F)$ of $\mathcal{I}$. Then, for any $U' << U$, the restrictions of f and g to U' may be regarded as morphisms $(U' \rightarrowtail U) \rightrightarrows (F' \rightarrowtail F)$ in $\mathcal{I}$; so by 4.4.10 they are coequalized by some $(F' \rightarrowtail F) \to (G' \rightarrowtail G)$ in $\mathcal{I}$; i.e. the equalizer of $U \rightrightarrows F' \to G'$ contains U'. Since U is covered by such opens U', it follows that f and g represent the same morphism $U \to \varinjlim_{\mathcal{I}} F'$; so the colimit has at most one section over U. But the opens U with $U \prec\!\prec X$ form a basis for $\mathcal{O}(X)$; so the colimit has at most one section over any open set.

The proof of initiality is generally similar. Given any ind-object $(A_j \mid j \in J)$ with colimit X, we have by definition a morphism from any admissible object F to some A_j. But if we had two such morphisms $F \rightrightarrows A_j$, and a subobject $F' << F$, then there must exist $j \to j'$ such that the composites $F' \rightarrowtail F \rightrightarrows A_j \to A_{j'}$ are equal; so for any such F' there is a unique equivalence class of morphisms $F' \to A_j$ which includes the restriction of some morphism defined on F. Taking this equivalence class for each object $(F' \rightarrowtail F)$ of $\mathcal{I}$, we obtain a morphism of ind-objects $(F' \mid (F' \rightarrowtail F) \in \mathcal{I}) \to (A_j \mid j \in J)$. However, given any object $(F' \rightarrowtail F)$ of $\mathcal{I}$, we can find F'' with $F' << F'' << F$ in $\mathrm{Sub}(F)$; then the inclusion $F' \rightarrowtail F''$ is a morphism $(F' \rightarrowtail F) \to (F'' \rightarrowtail F)$ in $\mathcal{I}$, and so if we are given any morphism of ind-objects with domain $(F' \mid (F' \rightarrowtail F) \in \mathcal{I})$, its $(F' \rightarrowtail F)$th component must contain some morphism which extends to F''. Hence by the argument already given there is exactly one morphism $(F' \mid (F' \rightarrowtail F) \in \mathcal{I}) \to (A_j \mid j \in J)$. □

Putting everything together, we now have

Theorem 4.4.12 *Let X be an internal locale in $\mathcal{S}$. Then $\mathbf{Sh}_{\mathcal{S}}(X)$ is exponentiable in $\mathfrak{BTop}/\mathcal{S}$ iff X is locally metastably compact.*

Proof One direction is 4.4.7. Conversely, if X is locally metastably compact, then by 4.4.11 Ind-$\mathbf{Sh}(X)_X$ has an initial object; and since local metastable compactness is a local condition, the same is true for all the fibres of $\varinjlim$: Ind-$\mathbf{Sh}(X) \to \mathbf{Sh}(X)$. So by 4.2.10 $\mathbf{Sh}(X)$ is a continuous category, and by 4.4.5 it is an exponentiable topos. □

We remark that, although we defined th
of arbitrary (diagrams of) sheaves on X,
purely lattice-theoretic condition on $\mathcal{O}(X)$,
we reduced the condition (E) for strong com
in 3.4.4, to a condition on a site of definition
which may be found in [542].

Suggestions for further reading: Johnst

e relation $U \prec\!\!\prec V$ in $\mathcal{O}(X)$ in terms
the definition can be reduced to a
using ideas similar to those by which
pactness of an $\mathcal{S}$-topos $\mathcal{E}$, introduced
for $\mathcal{E}$ (cf. 3.4.5). We omit the details,

one & Joyal [542], Niefield [892, 896].

C5

TOPOSES AS GROUPOIDS

C5.1 The descent theorems

The title of this chapter perhaps needs a few words of explanation. Our general thesis throughout Part C has been that the notion of (Grothendieck) topos is a generalization of the notion of topological space (or, better, of the notion of locale). In particular, toposes (or geometric morphisms between them) can be ascribed topological properties such as (local) compactness or (local) connectedness, as we saw in Chapter C3; and the theory of exponentiability for toposes, which we explored in Chapter C4, also has close parallels with that for spaces (or locales). Nevertheless, we know from the examples of toposes we have seen (in Section A2.1 and elsewhere) that, even if we restrict our attention to Grothendieck toposes, there are many which (in their origins, at least) do not appear at all 'spatial' in character. So one is tempted to ask whether this generalization is perhaps too large to be genuinely useful.

Our aim in this chapter is to demonstrate that the generalization from locales to (Grothendieck) toposes is not, after all, as broad as it appears. Specifically, we shall show that, by uniting the notion of locale with that of a groupoid, we obtain representation theorems enabling us to describe arbitrary Grothendieck toposes in terms of these two concepts. So the original informal picture of a Grothendieck topos, dating from the earliest days of work on étale cohomology of schemes, as 'a space whose points have enough internal structure to allow them to possess nontrivial automorphisms' is not as wide of the mark as all that.

There are several different versions of these representation theorems, and in the interests of comprehensiveness we shall prove more versions than any one reader may perhaps wish to see. But (with the exception of the 'Freyd representation' studied in Section C5.4) they all have a family resemblance: in particular, they depend on combining the idea that an arbitrary Grothendieck topos has a particular type of 'covering' by a localic topos, with theorems which say that the covering maps concerned are descent morphisms in $\mathfrak{BTop}$, in a suitable sense. The covering theorems will be proved in Section C5.2; in the present section we concern ourselves with the descent theorems.

Since we shall be making extensive use of pullbacks of geometric morphisms, we shall assume until further notice that all geometric morphisms under consideration are bounded. We shall also assume whenever necessary that the base

topos $\mathcal{S}$ over which we are working has a natural number object; and we shall use the term 'Grothendieck topos' loosely to mean 'bounded $\mathcal{S}$-topos', without explicitly specifying the base topos $\mathcal{S}$.

We begin by recalling the definition of descent morphisms from Sections B1.5 and B3.4. Given a bounded geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, we may form the diagram

$$\begin{tikzcd} \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \arrow[r, shift left=2, "\pi_{12}"] \arrow[r, "\pi_{13}"] \arrow[r, shift right=2, "\pi_{23}"'] & \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \arrow[r, shift left=2, "\pi_1"] \arrow[r, shift right=2, "\pi_2"'] & \mathcal{F} \arrow[l, "\Delta"'] \end{tikzcd}$$

(recall that we called such a diagram a *2-truncated simplicial topos* in Section B3.4 – we shall refer to this diagram as $\mathcal{F}_\bullet$), and we recall that *descent data* for an object B of $\mathcal{F}$ consists of a morphism $\theta\colon \pi_1^*(B) \to \pi_2^*(B)$ in $\mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ such that (modulo coherence isomorphisms) we have $\Delta^*(\theta) = 1_B$ and the diagram

$$\begin{tikzcd} \pi_1^*(B) \arrow[r, "\pi_{12}^*(\theta)"] \arrow[dr, "\pi_{13}^*(\theta)"'] & \pi_2^*(B) \arrow[d, "\pi_{23}^*(\theta)"] \\ & \pi_3^*(B) \end{tikzcd}$$

commutes in $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$. (Such a morphism θ is necessarily an isomorphism, as we saw in B1.5.1: the pullback of θ along the 'twist' isomorphism $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \to \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ is an inverse for θ.) If $B = f^*(A)$ for some $A \in \mathrm{ob}\, \mathcal{E}$, then the component at A of the canonical isomorphism $\pi_1^* f^* \cong \pi_2^* f^*$ constitutes descent data for B, so we can lift f^* to a functor $\mathcal{E} \to \mathbf{Desc}_f(\mathcal{F}_\bullet)$, where $\mathbf{Desc}_f(\mathcal{F}_\bullet)$ (or simply $\mathbf{Desc}(\mathcal{F}_\bullet)$, if we do not need to specify the morphism f explicitly) denotes the category of objects of $\mathcal{F}$ equipped with descent data. In fact, as we saw in B3.4.12, the category $\mathbf{Desc}(\mathcal{F}_\bullet)$ is always a bounded $\mathcal{E}$-topos, provided $\mathcal{E}$ has a natural number object, although for most of this and the next section we shall not need to assume this fact. (We shall need it in Section C5.3, however.)

We say that f is a *descent morphism* (resp. a *pre-descent morphism*) if the canonical functor $\mathcal{E} \to \mathbf{Desc}(\mathcal{F}_\bullet)$ is part of an equivalence (resp. full and faithful). (Given that this functor is the inverse image of a geometric morphism, the latter condition just says that $\mathbf{Desc}(\mathcal{F}_\bullet)$ is a connected $\mathcal{E}$-topos. Note that since $\mathcal{F} \to \mathbf{Desc}(\mathcal{F}_\bullet)$ is always surjective, a pre-descent morphism is necessarily a surjection.) If f is a descent morphism, we shall also say that *objects descend along* f. (This is not quite the sense in which we used the phrase 'f is a descent morphism' in Section B1.5; in the terminology of that section, our present condition is equivalent to saying that f is a descent morphism for the $\mathfrak{BTop}$-indexed category given by the forgetful functor $\mathfrak{BTop}^{\mathrm{op}} \to \mathfrak{Cat}$ which sends a topos to its underlying category, and a geometric morphism to its inverse image functor. In fact I. Moerdijk [835] has shown that open surjections *are* descent morphisms, in an appropriate 2-categorical sense, in $\mathfrak{BTop}$.)

We begin with the simplest case of the descent theorem, that in which f is a stable (right or left) Beck–Chevalley morphism in the sense of 2.4.16. We note that the morphism $\theta\colon \pi_1^*(B) \to \pi_2^*(B)$ may be transposed to a morphism $\overline{\theta}\colon B \to \pi_{1*}\pi_2^*(B)$ in $\mathcal{F}$, and if the pullback square of f against itself satisfies the Beck–Chevalley condition we may identify the codomain of $\overline{\theta}$ with $f^*f_*(B)$. Moreover, by the arguments in the proof of B1.5.5 (which use the Beck–Chevalley condition only for pullbacks of pullbacks of the particular morphism under consideration), saying that θ constitutes descent data for f is equivalent to saying that $\overline{\theta}$ is a coalgebra structure for the comonad induced by the adjunction $(f^* \dashv f_*)$. So we may immediately deduce the following result from A4.2.6, plus the results on stable Beck–Chevalley morphisms in 3.3.16 and 3.4.11:

Theorem 5.1.1 *Suppose $f\colon \mathcal{F} \to \mathcal{E}$ is bounded, surjective and either locally connected or tidy. Then objects descend along f.* □

In fact, however, this result is not best possible: both open and proper surjections are descent morphisms. To prove this, it is convenient to go by way of descent for locales, and then to regard objects as a special (discrete) case of locales: of course, we say that *locales descend along* $f\colon \mathcal{F} \to \mathcal{E}$ if the functor $\mathbf{Loc}(\mathcal{E}) \to \mathbf{Desc}(\mathbf{Loc}(\mathcal{F})_\bullet)$ induced by $f^*\colon \mathbf{Loc}(\mathcal{E}) \to \mathbf{Loc}(\mathcal{F})$ is part of an equivalence, where $\mathbf{Desc}(\mathbf{Loc}(\mathcal{F})_\bullet)$ denotes the category whose objects are internal locales Y in $\mathcal{F}$ equipped with morphisms $\theta\colon \pi_1^*(Y) \to \pi_2^*(Y)$ in $\mathbf{Loc}(\mathcal{F} \times_{\mathcal{E}} \mathcal{F})$ satisfying the usual unit and cocycle conditions.

Lemma 5.1.2 *Locales descend along hyperconnected morphisms.*

Proof In 2.4.11 we saw that, for any pullback square in $\mathfrak{BTop}$, the Beck–Chevalley condition holds for internal locales in the toposes concerned. Thus, by (the dual of) the argument of B1.5.5, we may always identify $\mathbf{Desc}_f(\mathbf{Loc}(\mathcal{F})_\bullet)$ with the category of algebras for the monad on $\mathbf{Loc}(\mathcal{F})$ induced by $f^*\colon \mathbf{Loc}(\mathcal{E}) \to \mathbf{Loc}(\mathcal{F})$ and its left adjoint $f_!$. The problem is thus to show that f^* is monadic at the level of locales. But for hyperconnected morphisms f this follows easily from 2.4.15, which says that the adjunction is a reflection in this case. □

Lemma 5.1.3 *Suppose given a commutative triangle of geometric morphisms*

$$\begin{tikzcd} \mathcal{G} \arrow[r, "h"] \arrow[dr, "g"'] & \mathcal{F} \arrow[d, "f"] \\ & \mathcal{E} \end{tikzcd}$$

where h is hyperconnected. Then locales descend along g iff they do so along f.

Proof First suppose locales descend along g, and let Y be an internal locale in $\mathcal{F}$ equipped with descent data θ for f. By applying the inverse image of

the geometric morphism $h \times h \colon \mathcal{G} \times_{\mathcal{E}} \mathcal{G} \to \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ to the descent data, one obtains descent data for h^*Y relative to g, so by assumption there exists a locale X in $\mathcal{E}$ satisfying $g^*(X) \cong h^*f^*(X) \cong h^*(Y)$. Since h^* is full and faithful on locales, we deduce $f^*X \cong Y$. The construction is functorial in (Y, θ), and clearly defines an inverse (up to isomorphism) for the canonical functor $\mathbf{Loc}(\mathcal{E}) \to \mathbf{Desc}_f(\mathbf{Loc}(\mathcal{F})_\bullet)$.

Conversely, suppose locales descend along f, and let Z be an internal locale in $\mathcal{G}$ equipped with descent data for g. Once again, we may pull the data back along $\mathcal{G} \times_{\mathcal{F}} \mathcal{G} \to \mathcal{G} \times_{\mathcal{E}} \mathcal{G}$, to obtain descent data for Z relative to h, and hence by 5.1.2 we obtain a locale Y in $\mathcal{F}$ such that $h^*Y \cong Z$. Now $h \times h \colon \mathcal{G} \times_{\mathcal{E}} \mathcal{G} \to \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ is hyperconnected, since it is a composite of pullbacks of h, so the isomorphism $h^*\pi_1^*(Y) \cong \pi_1^*(Z) \to \pi_2^*(Z) \cong h^*\pi_2^*(Y)$ 'descends' to an isomorphism $\pi_1^*(Y) \to \pi_2^*(Y)$ in $\mathbf{Loc}(\mathcal{F} \times_{\mathcal{E}} \mathcal{F})$, which is readily verified to be descent data for Y relative to f. Thus we obtain a locale X in $\mathcal{E}$ such that $f^*X \cong Y$ and hence $g^*X \cong h^*Y \cong Z$; it is again easy to verify that this construction yields an inverse (up to natural isomorphism) for the functor $\mathbf{Loc}(\mathcal{E}) \to \mathbf{Desc}_g(\mathbf{Loc}(\mathcal{G})_\bullet)$. $\square$

Proposition 5.1.4 *For any topos $\mathcal{S}$, open surjections and proper surjections of locales are descent morphisms (in the sense of* B1.5.5*) in* $\mathbf{Loc}(\mathcal{S})$.

Proof We use B1.5.7, which provided a criterion for a class of morphisms in a cartesian category to consist of descent morphisms. We already know that both open surjections and proper surjections satisfy the first three conditions of B1.5.7, by the results of 3.1.11 and 3.1.12 in the open case and by 3.2.6 in the proper case. So we need only verify condition (iv), which says that if we are given an equivalence relation $a, b \colon R \rightrightarrows X$ in $\mathbf{Loc}(\mathcal{S})$ with a and b open (respectively, proper), then the coequalizer $q \colon X \to Y$ of a and b is open (resp. proper) and pullback-stable.

The first assertion is easy in the open case, since by 1.5.3(iii) we know that open maps are precisely those for which the corresponding frame maps are homomorphisms of complete Heyting algebras; if a and b both have this property, then clearly the equalizer of a^* and b^* in $\mathbf{Frm}(\mathcal{S})$ is a sub-(complete Heyting algebra) of $\mathcal{O}(X)$. Alternatively, we could employ the dual of the following argument, which we need for the proper case.

Suppose a and b are proper, and consider the composite $a_*b^* \colon \mathcal{O}(X) \to \mathcal{O}(X)$. We note first that, since (a, b) is reflexive, we have $a_*b^* \leq a_*s_*s^*b^* = 1_{\mathcal{O}(X)}$ (where $s \colon X \to R$ is the diagonal map). Symmetry of (a, b) tells us that $a_*b^* = b_*a^*$; and transitivity tells us that it is idempotent, for if we form the pullback

$$\begin{tikzcd} R \times_X R \arrow[r, "d"] \arrow[d, "c"'] & R \arrow[d, "a"] \\ R \arrow[r, "b"] & X \end{tikzcd}$$

and let $t\colon R \times_X R \to R$ be the map expressing transitivity, we have $a_*b^*a_*b^* = a_*c_*d^*b^* = a_*t_*t^*b^* \geq a_*b^*$, using the Beck–Chevalley condition for the pullback square at the first step. Let $\mathcal{O}(Y)$ denote the set of fixed points of a_*b^*; then $\mathcal{O}(Y)$ is a subframe of $\mathcal{O}(X)$, since the inclusion $q^*\colon \mathcal{O}(Y) \to \mathcal{O}(X)$ has a right adjoint q_*, and it preserves finite meets since a_*b^* does. Also, the image of q^* is precisely the equalizer of a^* and b^*; for from $a^*(U) = b^*(U)$ we deduce $q^*q_*(U) = a_*b^*(U) = a_*a^*(U) \geq U$, whence $q_*q^*(U) = U$, and conversely if U is fixed by q^*q_* then $b^*(U) \geq a^*a_*b^*(U) = a^*q^*q_*(U) = a^*(U)$, and similarly $a^*(U) \geq b^*(U)$. Thus $q\colon X \to Y$ is the coequalizer of a and b in $\mathbf{Loc}(\mathcal{S})$.

To show that q is proper, we first note that q_* preserves directed joins, since $q^*q_* = a_*b^*$ does and q^* reflects joins. For the $\mathcal{O}(Y)$-linearity condition, suppose $U \in \mathcal{O}(X)$ and $V \in \mathcal{O}(Y)$: then

$$\begin{aligned} q^*q_*(U \cup q^*(V)) &= a_*b^*(U \cup q^*(V)) \\ &= a_*(b^*(U) \cup b^*q^*(V)) \quad \text{since } b^* \text{ preserves finite joins} \\ &= a_*(b^*(U) \cup a^*q^*(V)) \quad \text{since } qa = qb \\ &= a_*b^*(U) \cup q^*(V) \quad \text{since } a \text{ is proper} \\ &= q^*(q_*(U) \cup V), \end{aligned}$$

but q^* is injective, so $q_*(U \cup q^*(V)) = q_*(U) \cup V$.

To show that the coequalizer diagram $R \rightrightarrows X \to Y$ is preserved under pullback in $\mathbf{Loc}(\mathcal{S})$, note that the identities we have etablished, plus the fact that a and q are surjective, say that

$$\begin{tikzcd} \mathcal{O}(Y) \arrow[r, shift left, "q^*"] & \mathcal{O}(X) \arrow[l, shift left, "q_*"] \arrow[r, shift left=3, "b^*"] \arrow[r, shift left, "a^*"'] & \mathcal{O}(R) \arrow[l, shift left, "a_*"] \end{tikzcd}$$

is a split equalizer diagram of internal preframes in $\mathcal{S}$, and the maps involved are all $\mathcal{O}(Y)$-linear. Pulling back along a locale morphism $Y' \to Y$ corresponds to applying the functor $(-) \odot_{\mathcal{O}(Y)} \mathcal{O}(Y')$ from $\mathcal{O}(Y)$-modules to $\mathcal{O}(Y')$-modules in $\mathbf{PFrm}(\mathcal{S})$ (cf. 1.1.9); so it preserves this equalizer. And the forgetful functor from frames to preframes creates equalizers, so the diagram remains an equalizer in $\mathbf{Frm}(\mathcal{S})$, that is a coequalizer in $\mathbf{Loc}(\mathcal{S})$. The argument for this part of the proof in the 'open' case is similar, using complete join-semilattices in place of preframes. □

Putting together the last three results, we have

Theorem 5.1.5 *Locales descend along arbitrary open or proper surjections in* $\mathfrak{BTop}$.

Proof Let $f\colon \mathcal{F} \to \mathcal{E}$ be a surjection which is either open or proper. By 5.1.3, locales descend along f iff they descend along its localic part, which we may

write as $\mathbf{Sh}_{\mathcal{E}}(X) \to \mathcal{E}$ for an internal locale X in $\mathcal{E}$ which is positive and either open or compact as appropriate. But now internal locales in $\mathbf{Sh}_{\mathcal{E}}(X)$ correspond to objects of $\mathbf{Loc}(\mathcal{E})/X$, by 1.6.3; so we are reduced to proving that $X \to 1$ is a descent morphism in $\mathbf{Loc}(\mathcal{E})$. And this follows from 5.1.4. □

To deduce that objects descend along an open or proper surjection f, we use the identification of objects with discrete locales, and the characterization of the latter as those locales X such that $X \to 1$ and the diagonal $X \to X \times X$ are open (3.1.15). Since $f^*\colon \mathbf{Loc}(\mathcal{E}) \to \mathbf{Loc}(\mathcal{F})$ preserves finite products (being a right adjoint), it therefore suffices to know that openness of locale morphisms descends along f, i.e. if $g\colon X \to Y$ is a morphism in $\mathbf{Loc}(\mathcal{E})$ such that $f^*(g)$ is open in $\mathbf{Loc}(\mathcal{F})$, then g itself is open. But we proved this in 3.2.23; so we may now deduce

Theorem 5.1.6 *Open surjections, and proper surjections, are descent morphisms in $\mathfrak{BTop}$.* □

From the fact that both objects and locales descend along a morphism, we may deduce many other descent properties. In particular, most of the properties of geometric morphisms which we studied in Chapter C3 descend along either open or proper surjections:

Corollary 5.1.7 *Suppose given a pullback square*

$$\begin{tikzcd} \mathcal{H} \arrow[r, "k"] \arrow[d, "h"] & \mathcal{G} \arrow[d, "g"] \\ \mathcal{F} \arrow[r, "f"] & \mathcal{E} \end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$. where f is surjective and either open or proper. If h has any of the following properties: surjective, open, proper, hyperconnected, stably connected, locally connected, tidy, atomic, separated, local, totally connected, then g has the corresponding property.

Proof 'Surjective' is obvious, and 'open' and 'proper' were dealt with in 3.2.23.

For 'hyperconnected', let X be the internal locale in $\mathcal{E}$ corresponding to the localic part of the hyperconnected-localic factorization of g; then $f^*X \cong 1$ in $\mathbf{Loc}(\mathcal{F})$. But this isomorphism must descend along f, so $X \cong 1$ in $\mathbf{Loc}(\mathcal{E})$; i.e. g is hyperconnected.

If h is stably connected (that is, all pullbacks of h are connected), it suffices to prove that g is connected, since any pullback of g can be pulled back along an open surjection to a pullback of h. But we already know that g is surjective, so we need to prove that g^* is full. Suppose we have objects A and B of $\mathcal{E}$ and a morphism $a\colon g^*A \to g^*B$ in $\mathcal{G}$; then k^*a is a morphism

$h^*f^*A \cong k^*g^*A \to k^*g^*B \cong h^*f^*B$ in $\mathcal{H}$, so it derives from a unique morphism $b: f^*A \to f^*B$ in $\mathcal{F}$. Moreover, k^*a commutes with the descent data which k^*g^*A and k^*g^*B possess relative to the open or proper surjection k; so, since the pullback of h to the kernel-pair of f is still connected, we see that b must commute with the descent data on f^*A and f^*B relative to f. Hence it is of the form h^*c for a morphism $c: A \to B$ in $\mathcal{E}$; then since $k^*g^*c \cong h^*f^*c = h^*b \cong k^*a$ and k^* is faithful we deduce that $g^*c = a$.

For local connectedness, we use the criterion of 3.3.5(ii): we need to show that for any object B of $\mathcal{G}$ we can factor the composite $\mathcal{G}/B \to \mathcal{G} \to \mathcal{E}$ as a connected morphism followed by a local homeomorphism. In fact it suffices to consider the case $B = 1$, since local homeomorphisms pull back to local homeomorphisms. We know we have such a factorization $\mathcal{H} \to \mathcal{F}/C \to \mathcal{F}$ of h, where $C = h_!1$; now consider the diagram

$$\begin{tikzcd}
\mathcal{H} \times_{\mathcal{G}} \mathcal{H} \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] \arrow[d] & \mathcal{H} \arrow[d, "h"] \\
\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] & \mathcal{F}
\end{tikzcd} .$$

Here both the squares are pullbacks; so since the factorization of 3.3.5(i) is stable under pullback, by the Beck–Chevalley condition for locally connected morphisms, we have an isomorphism $\pi_1^*C \cong \pi_2^*C$ in $\mathcal{F} \times_{\mathcal{E}} \mathcal{F}$. It is easy to see that this isomorphism satisfies the unit and cocycle conditions for descent data; so we deduce that $C \cong f^*A$ for an object A of $\mathcal{E}$. Moreover, the factorization $h': \mathcal{H} \to \mathcal{F}/C$ of h through $\mathcal{F}/C \to \mathcal{F}$ corresponds by B3.2.8(b) to a morphism $1 \to h^*C \cong k^*g^*A$ in $\mathcal{H}$; it is again easy to see that this morphism commutes with the descent data on k^*g^*A, and so derives from a morphism $1 \to g^*A$ in $\mathcal{G}$ – equivalently, to a factorization of g through $\mathcal{E}/A \to \mathcal{E}$. Now we have a pullback square

$$\begin{tikzcd}
\mathcal{H} \arrow[r, "k"] \arrow[d, "h'"] & \mathcal{G} \arrow[d, "g'"] \\
\mathcal{F}/C \arrow[r, "f/A"] & \mathcal{E}/A
\end{tikzcd}$$

where h' is stably connected (since it is also locally connected); so by the previous argument we deduce that g' is connected.

For the tidy case, we use the Beck–Chevalley condition for the pullback squares

$$\begin{tikzcd} \mathcal{H}\times_{\mathcal{G}}\mathcal{H} \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] \arrow[d] & \mathcal{H} \arrow[d, "h"] \\ \mathcal{F}\times_{\mathcal{E}}\mathcal{F} \arrow[r, shift left, "\pi_1"] \arrow[r, shift right, "\pi_2"'] & \mathcal{F} \end{tikzcd}$$

to deduce that, for any object B of $\mathcal{G}$, the object h_*k^*B of $\mathcal{F}$ can be equipped with descent data relative to f, and so can be written as f^*qB for some object qB of $\mathcal{E}$. It is clear that $B \mapsto qB$ is functorial; also, the natural maps $f^*A \to h_*h^*f^*A \cong h_*k^*g^*A \cong f^*qg^*A$ and $k^*g^*qB \cong h^*f^*qB \cong h^*h_*k^*B \to k^*B$ descend along f and k respectively, to yield natural transformations $1_{\mathcal{E}} \to qg^*$ and $g^*q \to 1_{\mathcal{G}}$. And since f^* and k^* are faithful, these transformations satisfy the triangular identities for an adjunction $(g^* \dashv q)$, i.e. we may identify q with g_*. Thus we have shown that the square satisfies the Beck–Chevalley condition. It now follows that if $\mathbb{C}$ is any filtered internal category in (a slice of) $\mathcal{E}$, then f^*g_* preserves colimits of shape $\mathbb{C}$, since h_*k^* does; but f^* preserves colimits and is conservative, so g_* preserves colimits of shape $\mathbb{C}$.

For the atomic case, we use the characterization of atomic morphisms in 3.5.14: it is easy to see that the diagonal Δ_h is a pullback of Δ_g along a morphism which is itself a pullback of f, and hence an open or proper surjection. A similar argument deals with the separated case.

For the case when h is local, we first note that by cases already dealt with g must be connected and tidy; so the Beck–Chevalley condition $f^*g_* \cong h_*k^*$ holds. But f^* and k^* are comonadic, so we may use (the dual of) the adjoint lifting theorem A1.1.3(i) to lift the right adjoint of h_* to one for g_*. By 3.6.1(iv), this is sufficient.

Finally, suppose h is totally connected. Then g is connected and locally connected, by cases already dealt with, so we need only show that the left adjoint $g_!$ of g^* is cartesian. But $f^*g_! \cong h_!k^*$, by the Beck–Chevalley condition for locally connected morphisms; $h_!k^*$ is cartesian by assumption, and f^* is cartesian and conservative, so the result follows. □

Next, we mention a variant of the notion of descent morphism, namely that of a lax descent morphism. Suppose given a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$ between bounded $\mathcal{S}$-toposes; instead of forming the pullback of f against itself, we could

alternatively form the comma object

$$\begin{tikzcd}
(f\downarrow f) \arrow[r, "d_1"] \arrow[d, "d_0"'] \arrow[dr, phantom, "\downarrow"] & \mathcal{F} \arrow[d, "f"] \\
\mathcal{F} \arrow[r, "f"] & \mathcal{E}
\end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$, and then the pullback

$$\begin{tikzcd}
(f\downarrow f\downarrow f) \arrow[r] \arrow[d] & (f\downarrow f) \arrow[d, "d_0"] \\
(f\downarrow f) \arrow[r] & \mathcal{F}
\end{tikzcd} ,$$

and so on. ($(f\downarrow f\downarrow f)$ is the universal solution to the problem of finding an object equipped with three morphisms d_0, d_1, d_2 to $\mathcal{F}$ and two 2-cells $fd_2 \to fd_1 \to fd_0$.) It is easy to see that this gives rise to a simplicial object in $\mathfrak{BTop}/\mathcal{S}$ (in fact an internal category, though not an internal groupoid), and f^* induces a functor from $\mathcal{E}$ to the category of 'lax descent data' for (the 2-truncation of) this simplicial object. The notion of lax descent data was defined in B3.4.14(a), and we observed there that the category of lax descent data for a 2-truncated simplicial object in $\mathfrak{BTop}/\mathcal{S}$ is always a bounded $\mathcal{S}$-topos. Of course, we say that f is a *lax descent morphism* in $\mathfrak{BTop}/\mathcal{S}$ if this functor is part of an equivalence. (Note that this notion depends on the base topos over which f is considered as lying, since the comma object $(f\downarrow f)$, unlike the pullback $\mathcal{F}\times_{\mathcal{E}}\mathcal{F}$, changes if we vary $\mathcal{S}$. In particular, if we take $\mathcal{S}$ to be $\mathcal{E}$ itself then the comma object coincides with the pullback, so f is a lax descent morphism over $\mathcal{E}$ iff it is a descent morphism as previously defined.)

In 2.3.17 and 3.4.18, we noted that the Beck–Chevalley condition holds for comma squares in $\mathfrak{BTop}/\mathcal{S}$, one edge of which is either $\mathcal{S}$-essential or relatively tidy. From this, and from the fact that the 'opposite edges' of such squares are actually locally connected (resp. tidy), we may deduce

Proposition 5.1.8 *$\mathcal{S}$-essential surjections, and relatively tidy surjections, are lax descent morphisms in $\mathfrak{BTop}/\mathcal{S}$.* □

The proof is exactly like that of 5.1.1.

The reader will have observed that, in the proof of 5.1.4, we did not show that an equivalence relation $a, b\colon R \rightrightarrows X$ in $\mathbf{Loc}(\mathcal{S})$, for which a and b are either open or proper, is necessarily effective. Indeed, this is not true in general (and we deliberately omitted this condition, which would have simplified the proof of

B1.5.7, from the hypotheses of that scholium, in order that we could apply it in the situation of 5.1.4). However, it is useful to have a criterion for an equivalence relation whose 'legs' are open or proper to be effective in $\mathbf{Loc}(\mathcal{S})$.

Lemma 5.1.9 *Let $a, b\colon R \rightrightarrows X$ be an equivalence relation in $\mathbf{Loc}(\mathcal{S})$, and suppose either that a is open and $(a,b)\colon R \rightarrowtail X \times X$ is proper, or that a is proper and (a,b) is open. Then (a,b) is effective.*

Proof The two cases are, as usual, dual to each other; we shall prove the first. Let $q\colon X \twoheadrightarrow Y$ be the coequalizer of a and b, and $c, d\colon S \to X$ the kernel-pair of q; then $\mathcal{O}(S)$ may be identified with the tensor product $\mathcal{O}(X) \otimes_{\mathcal{O}(Y)} \mathcal{O}(X)$ of $\mathcal{O}(Y)$-modules in $\mathbf{CjSLat}(\mathcal{S})$, by 1.1.9. We claim that the frame homomorphism $\mathcal{O}(S) \to \mathcal{O}(R)$ corresponding to the factorization $u\colon R \to S$ of (a,b) through (c,d) reflects the bottom element; clearly, this homomorphism sends a generator $U \otimes V$ of the tensor product (where $U, V \in \mathcal{O}(X)$) to the element $a^*(U) \cap b^*(V)$ of $\mathcal{O}(R)$.

But if $a^*(U) \cap b^*(V) = \emptyset$, then we have

$$\begin{aligned} \emptyset &= a_!(a^*(U) \cap b^*(V)) && \text{since } a_! \text{ preserves joins} \\ &= U \cap a_! b^*(V) && \text{since } a \text{ is open} \\ &= U \cap q^* q_!(V) && \text{by the proof of 5.1.4.} \end{aligned}$$

But we also have $V \cap q^* q_!(V) = V$, so $U \cap V = \emptyset \cap V = \emptyset$ as an element of the tensor product $\mathcal{O}(X) \otimes_{\mathcal{O}(Y)} \mathcal{O}(X)$. Since every element of $\mathcal{O}(S)$ is a join of elements of the form $U \otimes V$, the claim is established. But if $R \rightarrowtail X \times X$ is proper, then it is a closed inclusion by 3.2.6; hence $R \to S$ is also a closed inclusion. And if it is the closed sublocale $\complement W$ for an open $W \in \mathcal{O}(S)$, then the foregoing argument shows that $W = \emptyset$. So $R \to S$ is an isomorphism.

The dual case is similar, using the preframe tensor product $\mathcal{O}(X) \odot_{\mathcal{O}(Y)} \mathcal{O}(X)$ and showing that the frame homomorphism from the latter to $\mathcal{O}(R)$ reflects the top element. □

Remark 5.1.10 The result of 5.1.9 is not best possible. By examining the proof given in the open case, we see that it actually shows that the frame homomorphism $u^*\colon \mathcal{O}(S) \to \mathcal{O}(R)$ has the property that if $u^*(W) \leq a^*(U)$, then $W \leq c^*(U)$; for if W is a union of elements $U_i \otimes V_i$, then for each i we obtain $U_i \cap q^* q_!(V_i) = a_!(a^* U_i \cap b^* V_i) \leq a_! a^* U \leq U$, and so we can rewrite $U_i \otimes V_i$ as $(U_i \cap q^* q_! V_i) \otimes V_i$, which lies below $(U \otimes 1) = c^*(U)$. So the argument in fact shows that the image of u is fibrewise dense over X, in the sense defined in 1.1.22; hence if we merely assume that $R \rightarrowtail X \times X$ is a fibrewise closed inclusion over X, then we may conclude that $R \rightrightarrows X$ is effective. (Even this result is not best possible; we shall see in 5.3.4 that the fibrewise closedness is automatic if we simply assume that $R \rightarrowtail X \times X$ is an inclusion, and not merely a monomorphism in $\mathbf{Loc}$.)

There is a similar 'fibrewise' weakening of the second hypothesis in the proper case, which will be found in [1203].

Before leaving this section, we prove a long-promised result, which enables us to extend our results on cofiltered limits of toposes from the countable case to the general case. The key lemma could in fact have been proved at any time since 3.1.20 (indeed, the major part of it could have been proved at any time since 1.2.8), but since its applications all involve the descent theorem it seemed sensible to place it here.

Lemma 5.1.11 *Let $\mathcal{S}$ be a topos with a natural number object, and $\mathbb{C}$ a filtered internal category in $\mathcal{S}$. Then there exists an open localic surjection $p\colon \mathcal{S}' \to \mathcal{S}$ and a final functor $\mathbb{N} \to p^*\mathbb{C}$ in $\mathcal{S}'$, where $\mathbb{N}$ denotes the poset obtained by equipping the natural number object with the order-relation of* A2.5.13.

Proof As in 1.2.8, the basic idea is to take $\mathcal{S}'$ to be the classifying topos for the propositional theory of final functors $\mathbb{N} \to \mathbb{C}$. To define the syntactic site of this theory, we take the poset $\mathbb{P}$ in $\mathcal{S}$ whose elements are functors $f\colon [p] \to \mathbb{C}$ from finite cardinals to $\mathbb{C}$ (where we identify a cardinal $[p]$ with the strict down-segment $\{0, 1, \ldots, p-1\}$ of p in $\mathbb{N}$, and order it as a sub-poset of $\mathbb{N}$), ordered by $f \leq g$ iff f extends g. We impose a coverage T on $\mathbb{P}$ by saying that, for each $f\colon [p] \to \mathbb{C}$ in $\mathbb{P}$ and each $c\colon C_0$, the set of $g \leq f$ for which there exists a morphism $c \to g(q-1)$ in $\mathbb{C}$ (where $[q]$ is the domain of g) is a covering sieve. It is easy to check that the set of all such sieves satisfies the stability condition (C) in the definition of a coverage; so we have a topos $\mathcal{S}' = \mathbf{Sh}_{\mathcal{S}}(\mathbb{P}, T) \simeq \mathbf{Sh}_{\mathcal{S}}(X)$, where X is the locale corresponding to the frame of T-ideals of $\mathbb{P}$. Moreover, $\mathbb{P}$ has a top element and the T-covering sieves are all inhabited (the latter because $\mathbb{C}$ is filtered), so by 3.1.20 $\mathcal{S}' \to \mathcal{S}$ is an open surjection.

Now, for any $\mathcal{S}$-topos $q\colon \mathcal{F} \to \mathcal{S}$, geometric morphisms $\mathcal{F} \to \mathcal{S}'$ over $\mathcal{S}$ correspond to flat cover-preserving functors $\mathbb{P} \to \mathbb{F}$; but any such functor G defines a functor $g\colon \mathbb{N} \to q^*\mathbb{C}$ in $\mathcal{F}$ (namely, the union in $\mathcal{F}$ of all those $f\colon [p] \to \mathbb{C}$ which are mapped by G to 1), which has the property that

$$(\forall c\colon C_0)(\exists n\colon N)(\exists \alpha\colon C_1)(\alpha\colon c \to g(n))$$

is valid in $\mathcal{F}$. Since $\mathbb{C}$ is filtered, this condition is sufficient for finality. □

In view of B2.6.13, it would have been sufficient to prove 5.1.11 for directed posets $\mathbb{C}$, but in fact it is no harder to prove the result for arbitrary filtered categories. Using 5.1.11, we may now extend all our results on limits of inverse sequences in $\mathfrak{BTop}/\mathcal{S}$ to arbitrary cofiltered $\mathcal{S}$-indexed limits. For example, we have

Corollary 5.1.12 *The localic reflection functor $\mathfrak{BTop}/\mathcal{S} \to \mathfrak{LTop}/\mathcal{S}$ preserves arbitrary cofiltered $\mathcal{S}$-indexed limits.*

Proof Suppose $\mathbb{C}$ is a cofiltered internal category in $\mathcal{S}$, and suppose we have a diagram of bounded $\mathcal{S}$-toposes $(\mathcal{E}_c \mid c : C_0)$ of shape $\mathbb{C}$. Let $\mathcal{F}_c$ denote the localic reflection of $\mathcal{E}_c$, for each c. Applying 5.1.11 to $\mathbb{C}^{\mathrm{op}}$, we obtain an open surjection $\mathcal{S}' \to \mathcal{S}$ and an initial functor $g \colon \mathbb{N}^{\mathrm{op}} \to \mathbb{C}$ in $\mathcal{S}'$. If we form the pullbacks $\mathcal{E}'_c = \mathcal{S}' \times_{\mathcal{S}} \mathcal{E}_c$ and $\mathcal{F}'_c = \mathcal{S}' \times_{\mathcal{S}} \mathcal{F}_c$, then by 2.4.12 we know that $\mathcal{F}'_c$ is the reflection of $\mathcal{E}'_c$ in $\mathfrak{LTop}/\mathcal{S}'$; also, since pullback preserves limits, we have $\varprojlim_{\mathbb{C}}(\mathcal{E}'_c) \simeq \mathcal{S}' \times_{\mathcal{S}} \varprojlim_{\mathbb{C}}(\mathcal{E}_c)$, and similarly for the $\mathcal{F}_c$. But we also have $\varprojlim_{\mathbb{C}}(\mathcal{E}'_c) \simeq \varprojlim_n(\mathcal{E}'_{g(n)})$, and hence by 2.5.12 we know that $\varprojlim_{\mathbb{C}}(\mathcal{F}'_c)$ is the reflection of $\varprojlim_{\mathbb{C}}(\mathcal{E}'_c)$ in $\mathfrak{LTop}/\mathcal{S}'$. Now we have a pullback square

$$\begin{tikzcd}
\varprojlim_{\mathbb{C}}(\mathcal{E}'_c) \arrow[r] \arrow[d] & \varprojlim_{\mathbb{C}}(\mathcal{E}_c) \arrow[d] \\
\varprojlim_{\mathbb{C}}(\mathcal{F}'_c) \arrow[r] & \varprojlim_{\mathbb{C}}(\mathcal{F}_c)
\end{tikzcd}$$

of which the top and bottom edges are open surjections and the left vertical edge is hyperconnected; so the right vertical edge is hyperconnected by 5.1.7. But $\varprojlim_{\mathbb{C}}(\mathcal{F}_c)$ is localic over $\mathcal{S}$, since $\mathfrak{LTop}/\mathcal{S}$ (being reflective in $\mathfrak{BTop}/\mathcal{S}$) is closed under arbitrary limits; so the result follows. $\square$

Corollary 5.1.13 *Suppose given a cofiltered $\mathcal{S}$-indexed diagram in $\mathfrak{BTop}/\mathcal{S}$, whose transition maps are all surjective (resp. open and surjective, proper, connected and locally connected, connected and atomic, local, totally connected). Then the legs of the colimiting cone satisfy the same condition.*

Proof In all cases except the proper one, these follow in the same way as 5.1.12 from the corresponding results for limits of inverse sequences (2.5.14, 3.1.22, 3.3.13, 3.5.11, 3.6.7(v), 3.6.18(iv)), plus the facts (proved in 5.1.7) that all the classes of geometric morphisms concerned descend along open surjections.

Since we did not prove the result for inverse sequences of proper maps, we must argue rather differently in that case. First we observe that, without loss of generality, we may assume that the base topos $\mathcal{S}$ is one of the vertices of our diagram; if not, we choose a particular vertex $\mathcal{E}_c$ and work in $\mathfrak{BTop}/\mathcal{E}_c$, replacing our original index category $\mathbb{C}$ by the slice category $\mathbb{C}/c$. Now we form the hyperconnected–localic factorization $\mathcal{E}_c \to \mathcal{L}_c \to \mathcal{S}$ of each $\mathcal{E}_c \to \mathcal{S}$; since hyperconnected maps are proper, it follows from 3.2.16 that the transition maps between the $\mathcal{L}_c$ are proper, and hence that the maps from the inverse limit $\mathcal{L}_\infty$ to the $\mathcal{L}_c$ are proper by 3.2.11. In particular, since $\mathcal{S}$ is one of the $\mathcal{L}_c$, $\mathcal{L}_\infty$ is compact over $\mathcal{S}$. By 5.1.12, it follows that the limit $\mathcal{E}_\infty$ of our original diagram is also compact over $\mathcal{S}$. But we saw that we could replace $\mathcal{S}$ by any of the vertices of the diagram, so this in fact means that the maps $\mathcal{E}_\infty \to \mathcal{E}_c$ are all proper. $\square$

Suggestions for further reading: Grothendieck [421], Joyal & Tierney [560], Kock [627], Moerdijk [827,835], Vermeulen [1203].

C5.2 Groupoid representations

In this section we shall prove the fundamental representation theorem of A. Joyal and M. Tierney [560] to the effect that every Grothendieck topos is equivalent to the topos of continuous actions of a localic groupoid. Here, and throughout this section, 'Grothendieck topos' is once again used loosely to mean 'bounded $\mathcal{S}$-topos', where $\mathcal{S}$ is any topos with a natural number object (the natural number object is essential: see 5.2.2 below); but we shall as usual treat $\mathcal{S}$ as if it were the classical category of sets, pausing only occasionally to point out any extra difficulties that may be encountered in 'internalizing' the arguments to a general topos.

The key idea behind the representation theorem is to combine the descent theorem of the previous section with a theorem asserting that there are 'enough localic $\mathcal{S}$-toposes' in the sense that every bounded $\mathcal{S}$-topos is an open surjective image of a localic one. We shall give several different proofs of this theorem, since each of them throws slightly different light on the structure of the 2-category $\mathfrak{BTop}/\mathcal{S}$. We begin with the first published proof of the theorem, which is essentially due to R. Diaconescu [282] (although the concept of open map did not appear explicitly in Diaconescu's work).

Theorem 5.2.1 *Let $\mathcal{S}$ be a topos with a natural number object. For any bounded geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$, there exists an open surjection $f\colon \mathcal{F} \to \mathcal{E}$ such that the composite pf is localic.*

Proof We consider first the case when $\mathcal{E}$ has the form $[\mathbb{C}, \mathcal{S}]$ for an internal category $\mathbb{C}$ in $\mathcal{S}$; in this case we shall construct a functor $d\colon \mathbb{P} \to \mathbb{C}$, where $\mathbb{P}$ is a poset, which is surjective on objects (so that the geometric morphism it induces is surjective) and satisfies the condition of 3.1.2 (so that the morphism it induces is open). This clearly suffices, since $[\mathbb{P}, \mathcal{S}]$ is localic over $\mathcal{S}$ if $\mathbb{P}$ is a poset (A4.6.2(d)). We shall argue as if $\mathcal{S}$ were the topos of sets, but it should be clear that the argument can be internalized in any topos with a natural number object.

We take the objects of $\mathbb{P}$ to be all finite composable strings of morphisms of $\mathbb{C}$, ordered by setting $\vec{a} = (a_1, \ldots, a_n) \leq \vec{b} = (b_1, \ldots, b_k)$ iff $\vec{a}$ is a terminal segment of $\vec{b}$, i.e. $n \leq k$ and $a_i = b_{i+k-n}$ for all i. The functor $d\colon \mathbb{P} \to \mathbb{C}$ sends $(a_1, \ldots, a_n)$ to the codomain of a_1, and an inequality $\vec{a} \leq \vec{b}$ to the morphism of $\mathbb{C}$ obtained by composing the morphisms which are in $\vec{b}$ but not in $\vec{a}$. It is clear that d is surjective on objects, and that it satisfies the condition of 3.1.2 (indeed, in the latter, we may take r and i to be identity morphisms: given a string $\vec{a}$ and $b\colon d(\vec{a}) \to V$, we define $\vec{c}$ to be the string obtained by adjoining b to the beginning of $\vec{a}$, and then we have $\vec{a} \leq \vec{c}$ and $d(\vec{a} \leq \vec{c}) = b$).

For the general case, we represent $\mathcal{E}$ as a subtopos of one of the form $[\mathbb{C}, \mathcal{S}]$, and then form the pullback

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[d] & \mathcal{E} \arrow[d] \\ {[\mathbb{P}, \mathcal{S}]} \arrow[r] & {[\mathbb{C}, \mathcal{S}]} \end{tikzcd}$$

where $\mathbb{P}$ is constructed as above. Then 3.1.26 ensures that the top edge of this diagram is an open surjection; and the left-hand side is an inclusion, so $\mathcal{F}$ is localic over $\mathcal{S}$. □

Remark 5.2.2 The assumption that the base topos $\mathcal{S}$ has a natural number object cannot be omitted from the hypotheses of 5.2.1: the result fails when $\mathcal{S}$ is the topos $\mathbf{Set}_f$ of finite sets. Indeed, let $\mathcal{C}$ be the finite category with two objects U and V and three non-identity morphisms $e: U \to U$, $r: U \to V$ and $i: V \to U$ satisfying $ri = 1_V$ and $ir = e = e^2$ (equivalently, $\mathcal{C}$ is the Cauchy-completion of the two-element monoid which is not a group). Then we claim that *there is no open surjection $\mathcal{F} \to [\mathcal{C}, \mathbf{Set}_f]$ such that $\mathcal{F}$ is localic over* $\mathbf{Set}_f$. For we know by 2.2.22 that every bounded $\mathbf{Set}_f$-topos is of the form $[\mathcal{P}, \mathbf{Set}_f]$ where $\mathcal{P}$ is a finite category, and clearly the localic ones are those for which we can take $\mathcal{P}$ to be a finite poset. Moreover, since $\mathcal{C}$ is Cauchy-complete, we know by 2.2.23 that every geometric morphism $[\mathcal{P}, \mathbf{Set}_f] \to [\mathcal{C}, \mathbf{Set}_f]$ is induced by a functor $\mathcal{P} \to \mathcal{C}$. Thus finding a geometric morphism as above is equivalent to finding a finite poset $\mathcal{P}$ and a functor $f: \mathcal{P} \to \mathcal{C}$ whose image contains the object U (for otherwise the corresponding geometric morphism would factor through the inclusion $\mathbf{Set}_f \to [\mathcal{C}, \mathbf{Set}_f]$ whose inverse image is given by evaluation at V), and which satisfies the condition of 3.1.2. Suppose given such an f: let P_0 be an element of $\mathcal{P}$ with $f(P_0) = U$, and apply the condition of 3.1.2 to the morphism $e: f(P_0) \to U$. Since U is not a retract of any other object of $\mathcal{C}$, we deduce that there must exist $P_1 \geq P_0$ in $\mathcal{P}$ such that $f(P_1) = U$ and the inequality $P_0 \leq P_1$ maps to the endomorphism e. Clearly, $P_1 \neq P_0$, since e is not an identity morphism; so on repeating the argument we deduce that $\mathcal{P}$ contains an infinite ascending chain $P_0 < P_1 < P_2 < \cdots$; in particular, $\mathcal{P}$ is infinite, contradicting the assumption that $[\mathcal{P}, \mathbf{Set}_f]$ is a bounded $\mathbf{Set}_f$-topos. (A similar argument using condition (ii) of 3.2.14 shows that $[\mathcal{C}, \mathbf{Set}_f]$ does not admit any proper surjection from a localic $\mathbf{Set}_f$-topos.)

Remark 5.2.3 The functor $d: \mathbb{P} \to \mathbb{C}$ of 5.2.1 has a natural factorization through the counit map $\mathbb{F}C \to \mathbb{C}$, where $\mathbb{F}C$ is the free category on the underlying directed graph of $\mathbb{C}$ (that is, the category whose objects are those of $\mathbb{C}$ and whose morphisms are composable strings of morphisms of $\mathbb{C}$). Specifically, the factorization $\bar{d}$ sends an object $\vec{a}$ to the codomain of the first morphism in $\vec{a}$.

and an inequality $\vec{a} \leq \vec{b}$ to the string of morphisms in $\vec{a}$ but not in $\vec{b}$. It is clear that $\overline{d}$ is a (surjective) discrete opfibration, so that we may identify $[\mathbb{P}, \mathcal{S}]$ with a slice category of the form $[\mathbb{F}C, \mathcal{S}]/D$ where D is well-supported; moreover, the counit map $\mathbb{F}C \to \mathbb{C}$ is full and surjective on objects, so that by A4.6.9 the geometric morphism which it induces is hyperconnected. Since both hyperconnected morphisms and surjective local homeomorphisms are stable under pullback, we deduce that *for every bounded* $\mathcal{S}$*-topos* $\mathcal{E}$*, there is a hyperconnected morphism* $\mathcal{F} \to \mathcal{E}$ *and a well-supported object* D *of* $\mathcal{F}$ *such that* $\mathcal{F}/D$ *is localic over* $\mathcal{S}$.

The French term *étendue* is often used for a 'locally localic' topos, i.e. a topos $\mathcal{E}$ containing a well-supported object A such that $\mathcal{E}/A$ is localic (cf. [1051]). We note in passing:

Lemma 5.2.4 *Let* $\mathcal{C}$ *be a small category. Then the functor category* $[\mathcal{C}, \mathbf{Set}]$ *is an étendue iff every morphism of* $\mathcal{C}$ *is epic.*

Proof Since any slice of $[\mathcal{C}, \mathbf{Set}]$ is equivalent to a category of the same form (A1.1.7), it is easy to see that $[\mathcal{C}, \mathbf{Set}]$ is an étendue iff there exists a surjective discrete opfibration $\mathcal{P} \to \mathcal{C}$ such that $\mathcal{P}$ is a poset. But every morphism in a poset is epic, and it is easy to see that this property 'descends' along surjective discrete opfibrations. Conversely, if every morphism of $\mathcal{C}$ is epic, then the domains of the discrete opfibrations correponding to the functors $\mathcal{C}(U, -)$ are all posets; and their coproduct maps surjectively to $\mathcal{C}$. □

Of course, 5.2.4 may be 'internalized' to an arbitrary topos $\mathcal{S}$. Clearly, the free category on a directed graph has the property that all its morphisms are epic (and monic), so this provides an alternative proof of 5.2.3 (and hence of 5.2.1).

For completeness, we also note:

Lemma 5.2.5 *A Grothendieck topos* $\mathcal{E}$ *is an étendue iff there exists a site of definition* $(\mathcal{C}, J)$ *for* $\mathcal{E}$ *such that every morphism of* $\mathcal{C}$ *is monic.*

Proof One direction follows easily from 5.2.4: if we are given such a site $(\mathcal{C}, J)$, then there exists a well-supported object A of $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ such that $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]/A$ is localic over $\mathbf{Set}$, and hence $\mathbf{Sh}(\mathcal{C}, J)/i^*A$ is localic, where i is the inclusion $\mathbf{Sh}(\mathcal{C}, J) \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$.

Conversely, suppose $\mathcal{E}$ is an étendue; let A be a well-supported object of $\mathcal{E}$ such that $\mathcal{E}/A$ is localic over $\mathbf{Set}$. Let $\mathcal{C}$ be the (non-full!) subcategory of objects of $\mathcal{C}$ which admit monomorphisms to A, and monomorphisms between them; we claim that the inclusion $\mathcal{C} \to \mathcal{E}$ satisfies the conditions of 2.2.1 (for the canonical coverage on $\mathcal{E}$), so that by 2.2.16(i) $\mathcal{C}$ becomes a site of definition for $\mathcal{E}$ when equipped with the induced coverage. For the first condition, suppose given $f, g \colon B \rightrightarrows C$ in $\mathcal{E}$ with $f \neq g$; then $A^*f \neq A^*g$ in $\mathcal{E}/A$, and since the latter topos is localic there exists $A' \rightarrowtail A$ and a morphism $h \colon A' \to B$ such that $fh \neq gh$. So, for any object B of $\mathcal{E}$, the family of all morphisms from objects of $\mathcal{C}$ to B is epimorphic. For the second condition, suppose given $f \colon B \to A'$ in $\mathcal{E}$ where A' admits a monomorphism $m \colon A' \rightarrowtail A$; then, regarding the composite

mf as an object of $\mathcal{E}/A$, we have an epimorphic family of morphisms to it from subterminal objects – which says that we have an epimorphic family of morphisms $A'' \to B$ such that the composites $A'' \to B \to A'$ are monic, and so lie in the subcategory $\mathcal{C}$. □

The following refinement of 5.2.1 is sometimes of use.

Scholium 5.2.6 *Let $\mathcal{E}$ be a bounded $\mathcal{S}$-topos, and suppose $\mathcal{E}$ has a point $s: \mathcal{S} \to \mathcal{E}$ (that is, a section of the structure map $p: \mathcal{E} \to \mathcal{S}$). Then it is possible to choose the open surjection f of* 5.2.1 *so that s factors through it.*

Proof As before, we first consider the case when $\mathcal{E} = [\mathbb{C}, \mathcal{S}]$ for some internal category $\mathbb{C}$ in $\mathcal{S}$. Then we know that points of $\mathcal{E}$ correspond to $\mathbb{C}$-torsors in $\mathcal{S}$, by B3.2.7; equivalently, they are objects of the ind-completion of $\mathbb{C}$ (*qua* $\mathcal{S}$-indexed category), as defined in Section C4.2. So we may think of s as being defined by an ind-object $(U_\imath \mid i \in I)$ of $\mathbb{C}$; and, by B2.6.13, we may assume that the indexing category $\mathbb{I}$ is actually a directed poset. Now we modify the definition of the poset $\mathbb{P}$ constructed in 5.2.1: in addition to the objects considered there, we also take all the objects of $\mathbb{I}$ together with all pairs $(\vec{a}, i)$, where i is an object of $\mathbb{I}$ and $\vec{a}$ is a composable string of morphisms of $\mathbb{C}$ such that the domain of its last member is U_i. The ordering is defined by saying that the 'pure strings' are ordered as they were before, we have $(\vec{a}, i) \leq (\vec{b}, j)$ iff $i = j$ and $\vec{a}$ is a terminal segment of $\vec{b}$, and that $i \leq (\vec{a}, j)$ and $i \leq j$ (where i, j are both objects of $\mathbb{I}$) iff the relation $i \leq j$ holds in $\mathbb{I}$. The functor $d: \mathbb{P} \to \mathbb{C}$ sends a 'pure string' $\vec{a}$ or an object of the form $(\vec{a}, i)$ to the codomain of the first member of $\vec{a}$; it sends an object i of $\mathbb{I}$ to U_i, and an inequality $i \leq j$ to the appropriate morphism in the ind-object $(U_i \mid i \in I)$. The verification that d is surjective on objects, and that it satisfies the condition of 3.1.2 for openness, is exactly as in 5.2.1; but the inclusion $\mathbb{I} \to \mathbb{P}$ now defines an ind-object of $\mathbb{P}$, and hence a point of $[\mathbb{P}, \mathcal{S}]$, whose image under the geometric morphism induced by d is clearly the given point of $\mathcal{E}$.

In general, we have an inclusion $\mathcal{E} \to [\mathbb{C}, \mathcal{S}]$ for some $\mathbb{C}$; we choose $\mathbb{P}$ as above for the point $\tilde{s}$ of $[\mathbb{C}, \mathcal{S}]$ which is the image of our given point of $\mathcal{E}$, and form the pullback

$$\begin{tikzcd} \mathcal{F} \arrow[r] \arrow[d] & \mathcal{E} \arrow[d] \\ {[\mathbb{P}, \mathcal{S}]} \arrow[r] & {[\mathbb{C}, \mathcal{S}]} \end{tikzcd} .$$

Since the point $\tilde{s}$ factors through both $\mathcal{E}$ and $[\mathbb{P}, \mathcal{S}]$, we obtain a factorization of it through the pullback, as required. □

The argument of 5.2.6 may clearly be generalized from a single point to a whole $\mathcal{S}$-indexed family of points (that is, a morphism $\mathcal{S}/I \to \mathcal{E}$ for some I).

In this connection we should mention that C. Butz and I. Moerdijk [214] have shown that, if the topos $\mathcal{E}$ has enough points (cf. 2.2.12), then we may choose its open cover $\mathcal{F}$ to have enough points also; that is, we may take it to be $\mathbf{Sh}(X)$ for a space X, and not merely a locale.

We remark also that if, in the first part of the proof of 5.2.6, the chosen point of $[\mathbb{C}, \mathcal{S}]$ had happened to be defined by an ind-object $(U_n \mid n \in \mathbb{N})$ indexed by the directed set $\mathbb{N}$ of natural numbers, then we should not have needed to enlarge our original poset $\mathbb{P}$ at all, since it would already contain a 'copy' of this ind-object; but more complicated directed sets cannot be embedded in our original $\mathbb{P}$, and so we do have to enlarge it as indicated.

The result of 5.2.1 is not best possible: we may require stronger conditions on the morphism f.

Theorem 5.2.7 *Let $\mathcal{S}$ be a topos with a natural number object. For any bounded geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$, there exists a connected and locally connected morphism $f\colon \mathcal{F} \to \mathcal{E}$ such that the composite pf is localic.*

Proof Recall that in 3.3.8 we gave conditions on a functor between small categories for it to induce a connected and/or locally connected morphism of **Set**-valued functor categories; the arguments used there were constructive, and so can be applied to functors between internal categories in a topos $\mathcal{S}$. It is not possible, for a general internal category $\mathbb{C}$, to find a functor $f\colon \mathbb{P} \to \mathbb{C}$ satisfying these conditions such that $\mathbb{P}$ is a poset; but we saw in 3.3.9 that we can do so in the case when $\mathbb{C}$ is the particular internal category $\mathbb{S}_f$ of finite cardinals in $\mathcal{S}$. And this is enough; for the functor category $[\mathbb{S}_f, \mathcal{S}]$ is the object classifier $\mathcal{S}[\mathbb{O}]$ for $\mathfrak{Top}/\mathcal{S}$, as we saw in B3.2.9, and in D3.2.5 we show that every bounded $\mathcal{S}$-topos $\mathcal{E}$ admits a localic morphism to $\mathcal{S}[\mathbb{O}]$ (in fact it suffices to take the classifying morphism of any bound for $\mathcal{E}$ over $\mathcal{S}$). Thus we may form the pullback

$$\begin{tikzcd} \mathcal{F} \arrow[r, "h"] \arrow[d, "g"'] & {[\mathbb{P}, \mathcal{S}]} \arrow[d, "d"] \\ \mathcal{E} \arrow[r, "e"] & \mathcal{S}[\mathbb{O}] \end{tikzcd}$$

where e is localic and $\mathbb{P}$ is the poset of finite partial equivalence relations on N, as in 3.3.9; then g is connected and locally connected since such morphisms are stable under pullback by 3.3.15, and h is localic for a similar reason, so that $\mathcal{F}$ is localic over $\mathcal{S}$ since composites of localic morphisms are localic. □

Remarks 5.2.8 (a) Theorem 5.2.1 could have been proved in the same style as 5.2.7, by first constructing a covering of the particular internal category $\mathbb{S}_f$ satisfying the conditions of A4.2.7(b) and 3.1.2, and then pulling back along a localic morphism $\mathcal{E} \to [\mathbb{S}_f, \mathcal{S}]$. However, in that case it is no more difficult to construct the covering for an arbitrary internal category.

(*b*) We note that 5.2.7 is 'best possible' in the sense that we cannot impose any stronger conditions on the covering map $f: \mathcal{F} \to \mathcal{E}$, at least among those considered in Chapter C3. For if we require f to be atomic and connected, then it is hyperconnected by 3.5.4(i); but it is also localic since $\mathcal{F}$ is localic over $\mathcal{S}$, and hence an equivalence, so $\mathcal{E}$ itself must be localic over $\mathcal{S}$. Even if we were to weaken the condition of connectedness to surjectivity, we still could not require f to be atomic; for if it were, it would then be a local homeomorphism by 3.5.4(iii), and so $\mathcal{E}$ would be an étendue. (We saw in 5.2.3 that any Grothendieck topos $\mathcal{E}$ admits a hyperconnected covering by an étendue $\mathcal{F}$, but once again we cannot require this covering to be atomic; for if it were, then the composite $\mathcal{F}/A \to \mathcal{F} \to \mathcal{E}$ (where A is a well-supported object such that $\mathcal{F}/A$ is localic over $\mathcal{S}$) would be atomic and surjective. In Section C5.4 below we shall investigate the 'best possible' conditions we can impose on a class of toposes $\mathcal{F}$ such that every Grothendieck topos admits a connected atomic covering by a topos $\mathcal{F}$ in the class.) Nor can we require f to be totally connected in the sense of 3.6.16; for if it were, then its right adjoint $d: \mathcal{E} \to \mathcal{F}$ would be an inclusion, and this would again force $\mathcal{E}$ to be localic over $\mathcal{S}$. We shall also see in 5.2.10 below that we cannot in general require the covering map f to be proper, even if we abandon the openness requirement altogether.

(*c*) It is possible to give proofs of both 5.2.1 and 5.2.7 making more extensive use of the theory of classifying toposes. (Our reason for giving more prominence to the proofs above was mainly to make the results more accessible to any readers who have not (yet) familiarized themselves with the ideas of Part D; there is a sense in which the argument via geometric theories is the 'right' way to prove these results.) For 5.2.7, we note that if $\mathbb{P}$ is the poset of finite partial equivalence relations on N, then $[\mathbb{P}, \mathcal{S}]$ is none other than the classifying topos for the propositional theory of partial equivalence relations on N (since the latter is a cartesian propositional theory; over **Set**, it may be axiomatized by an $(\mathbb{N} \times \mathbb{N})$-indexed family of primitive propositions $R(n,n')$ with axioms $(R(n,n') \vdash R(n',n))$ and $((R(n,n') \wedge R(n',n'')) \vdash R(n,n''))$ for all n, n', n''). And the geometric morphism $[\mathbb{P}, \mathcal{S}] \to [\mathbb{S}_f, \mathcal{S}]$ induced by d is none other than the classifying map of the 'generic subquotient of N' which exists in this classifying topos. Thus, when viewed as a topos defined over $[\mathbb{S}_f, \mathcal{S}]$, it becomes the classifying topos for the theory of partial enumerations of a particular object G (that is, the theory of partial maps $N \rightharpoonup G$ whose image is the whole of G), namely the generic object itself. We encountered this theory in 3.3.11(*c*), where we saw that it gave rise to a connected and locally connected site. (Note also that we may proceed directly to the construction of the topos $\mathcal{F}$ appearing as a pullback in the proof of 5.2.7, as the classifying topos over $\mathcal{E}$ for the theory of partial enumerations of some bound B for $\mathcal{E}$ over $\mathcal{S}$.)

The construction which we gave in 5.2.1 for an open surjection from a localic topos to a general bounded $\mathcal{S}$-topos does not lend itself so easily to a description in terms of classifying toposes, but a closely related construction (which was in fact the one given by Joyal and Tierney [560]) may be obtained as follows. First

of all, we note a slight modification of the result of D3.2.5, which we quoted in the proof of 5.2.7 above: namely, that every Grothendieck topos $\mathcal{E}$ admits a localic morphism to the classifying topos $\mathcal{S}[\mathbb{O}_1]$ for the theory of inhabited objects (that is, the theory with one sort, no primitive symbols except equality, and the single axiom $(\top \vdash (\exists x)\top)$). We remark that $\mathbf{Set}[\mathbb{O}_1]$ may readily be identified as the functor category $[\mathcal{C}, \mathbf{Set}]$ where $\mathcal{C}$ is the category of nonempty finite sets (and all functions between them); but we shall not need this information. To obtain a localic morphism $\mathcal{E} \to \mathcal{S}[\mathbb{O}_1]$, it suffices to observe that any bound for $\mathcal{E}$ over $\mathcal{S}$ is necessarily well-supported, and so its classifying map factors through $\mathcal{S}[\mathbb{O}_1] \to \mathcal{S}[\mathbb{O}]$.

Thus it suffices to construct an open surjection $f: \mathcal{F} \to \mathcal{S}[\mathbb{O}_1]$ with $\mathcal{F}$ localic over $\mathcal{S}$: for this, we take $\mathcal{F}$ to be the classifying topos for the propositional theory of (total) equivalence relations on the natural number object N, and f to be the classifying map of the quotient of N by the generic equivalence relation. Then, viewed as a topos over $\mathcal{S}[\mathbb{O}_1]$, f classifies the theory of (total) enumerations of the generic inhabited object G_1; more generally, its pullback along a morphism $\mathcal{E} \to \mathcal{S}[\mathbb{O}_1]$ classifies the theory of enumerations of the particular object of $\mathcal{E}$ classified by the latter. But we saw in 1.2.8 above how to construct a site for the theory of enumerations of a given inhabited object A; we observed there that the locale to which it gives rise is a dense sublocale of a nonempty spatial locale, but it is actually strongly dense, since all covers in the site are inhabited (cf. the remark after 3.1.18). So the fact that it gives rise to an open surjection follows from 1.2.14(ii) and 3.1.14(ii).

We note that this construction, although rather simpler than the earlier one using partial enumerations, is actually substantially different from it: the internal locale in $\mathcal{E}$ which it produces is very far from being locally connected. Indeed, if the inhabited object A whose enumerations are being classified is decidable, then the locale is actually zero-dimensional, since each sub-basic open set (corresponding to a primitive proposition $p(n,a)$ of the theory of 1.2.8) has an open complement, corresponding to the disjunction of the $p(n,b)$ with $b \neq a$.

In the previous section, we proved the descent theorem for proper surjections as well as open ones. As we remarked above, not every Grothendieck topos admits a proper surjection from a localic topos; but 3.2.15 enables us to identify an important class of toposes which do.

Lemma 5.2.9 *Let $\mathcal{E}$ be a Grothendieck topos containing a pre-bound (in the sense of* D3.2.3*) which is Kuratowski-finite in the sense of* D5.4.1. *Then there is a proper surjection $f: \mathcal{F} \to \mathcal{E}$ such that $\mathcal{F}$ is localic over* **Set**.

Proof In D3.2.10, we show that the classifying topos for K-finite objects in **Set**-toposes may be taken to be the functor category $[\mathbf{Set}_{fe}, \mathbf{Set}]$, where $\mathbf{Set}_{fe}$ is the category of finite sets and surjections between them. By D3.2.6(i), a K-finite pre-bound for $\mathcal{E}$ corresponds to a localic geometric morphism $\mathcal{E} \to [\mathbf{Set}_{fe}, \mathbf{Set}]$. In 3.2.15, we constructed a proper surjection $[\mathcal{C}, \mathbf{Set}] \to [\mathbf{Set}_{fe}, \mathbf{Set}]$ where $\mathcal{C}$ is

a poset (it is surjective because the functor inducing it is surjective on objects); pulling this back along $\mathcal{E} \to [\mathbf{Set}_{fe}, \mathbf{Set}]$, we obtain the required localic cover of $\mathcal{E}$. □

The condition that $\mathcal{E}$ should have a K-finite pre-bound is equivalent to saying that it is the classifying topos for a (single-sorted) geometric theory whose models are all K-finite (cf. D3.2.5). The converse of 5.2.9 is false; we shall see a counterexample in 5.2.12 below. Nevertheless, not every Grothendieck topos admits a proper surjection from a localic topos:

Example 5.2.10 Let G be a (discrete) group. We claim that *the topos* $[G, \mathbf{Set}]$ *admits a proper surjection from a localic topos iff* G *is finite.* For we know that $[G, \mathbf{Set}]$ is a classifying topos for G-torsors (cf. B3.2.4(b)); hence a geometric morphism $\mathbf{Sh}(X) \to [G, \mathbf{Set}]$ corresponds to a local homeomorphism $p \colon E \to X$ which is locally isomorphic to $X \times G$. Now the slice category $[G, \mathbf{Set}]/G$ (where G, as an object of $[G, \mathbf{Set}]$, is the generic G-torsor) is equivalent to $\mathbf{Set}$; so if $\mathbf{Sh}(X) \to [G, \mathbf{Set}]$ were proper, its pullback along $[G, \mathbf{Set}]/G \to [G, \mathbf{Set}]$ would be a proper map $\mathbf{Sh}(E) \to \mathbf{Set}$, i.e. E would be compact. But E cannot be compact if G is infinite: for we can cover X by opens U_i (indexed by some set I) such that each $p^{-1}(U_i)$ is homeomorphic to $U_i \times G$, and then cover E by (the homeomorphic images of) the sets $U_i \times \{g\}$, $g \in G, i \in I$; and this cover clearly has no finite subcover. On the other hand, if G is finite, then every G-torsor is K-finite; so $[G, \mathbf{Set}]$ admits a proper surjection from a localic topos, by 5.2.9. (In fact the morphism $\mathbf{Set} \simeq [G, \mathbf{Set}]/G \to [G, \mathbf{Set}]$ is proper in this case.)

Now suppose $\mathcal{E}$ is a bounded $\mathcal{S}$-topos, and we have found a descent morphism $f \colon \mathcal{F} \to \mathcal{E}$ to $\mathcal{E}$ from a topos $\mathcal{F}$ which is localic over $\mathcal{S}$. Then f is itself localic by A4.6.2(f), and hence so are the projections $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \rightrightarrows \mathcal{F}$, by B3.3.6. Hence $\mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ is itself localic over $\mathcal{S}$, by A4.6.2(e). Similarly, the triple pullback $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \times_{\mathcal{E}} \mathcal{F}$ is localic over $\mathcal{S}$; so we can regard the entire (truncated) simplicial topos

$$\begin{tikzcd} \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \arrow[r, shift left=2] \arrow[r] \arrow[r, shift right=2] & \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \arrow[r, shift left=2] \arrow[r, leftarrow] \arrow[r, shift right=2] & \mathcal{F} \end{tikzcd}$$

as living in the category $\mathbf{Loc}(\mathcal{S})$ of internal locales in $\mathcal{S}$. Moreover, the inclusion $\mathbf{Loc}(\mathcal{S}) \to \mathfrak{Top}/\mathcal{S}$ preserves limits, by 1.4.8; hence the simplicial locale $G_\bullet$ defined by setting the $(n+1)$-fold pullback of $\mathcal{F}$ over $\mathcal{E}$ equal to $\mathbf{Sh}_{\mathcal{S}}(G_n)$ inherits the same limit-preservation properties as the simplicial topos. In particular, the diagram

$$\begin{tikzcd} G_2 \arrow[r, "d_0"] \arrow[d, "d_2"] & G_1 \arrow[d, "d_1"] \\ G_1 \arrow[r, "d_0"] & G_0 \end{tikzcd}$$

is a pullback in $\mathbf{Loc}(\mathcal{S})$; but this (plus preservation of the pullbacks involving G_3, to ensure associativity of the composition defined by d_1^2) means that $G_\bullet$ is an internal category in $\mathbf{Loc}(\mathcal{S})$, as we saw in B2.3.1. And since we have a twist isomorphism $G_1 \to G_1$ satisfying the appropriate equations, it is actually an internal groupoid in $\mathbf{Loc}(\mathcal{S})$.

As we observed in B3.4.14(*b*), the functoriality of the assignment $X \mapsto \mathbf{Sh}_{\mathcal{S}}(X)$ is all we need to conclude that an internal groupoid $\mathbb{G}$ in $\mathbf{Loc}(\mathcal{S})$ gives rise to a topos $\mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$ of continuous actions of $\mathbb{G}$: the objects of this topos are sheaves on G_0 (equivalently, local homeomorphisms $E \to G_0$) equipped with action maps $d_1^*(E) \to d_0^*(E)$ over G_1 satisfying the unit and cocycle conditions. However, in the case currently under consideration, we do not need the proof in B3.4.14 that this category is a topos, since the descent theorem tells us that we have $\mathbf{Cont}_{\mathcal{S}}(\mathbb{G}) \simeq \mathbf{Desc}(\mathcal{F}_\bullet) \simeq \mathcal{E}$.

We shall say that a localic groupoid $\mathbb{G}$ is *open* (resp. *proper, locally connected, tidy, ...*) if its domain and codomain maps $d_1, d_0 \colon G_1 \rightrightarrows G_0$ have the property in question. Note that this will automatically be the case (for any pullback-stable property of geometric morphisms) if the covering morphism $\mathcal{F} \to \mathcal{E}$ from which we start has the corresponding property; also that d_0 has one of the properties iff d_1 does, since the inverse map $G_1 \to G_1$ is an isomorphism $d_0 \cong d_1$ in $\mathbf{Loc}(\mathcal{S})/G_0$. Thus the following results are all now immediate from the covering theorems proved earlier in this section, plus the descent theorems of Section C5.1.

Theorem 5.2.11 *Let $\mathcal{S}$ be a topos with a natural number object.*

(i) *For any bounded $\mathcal{S}$-topos $\mathcal{E}$, there exists a connected and locally connected localic groupoid $\mathbb{G}$ in $\mathcal{S}$ such that $\mathcal{E} \simeq \mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$.*

(ii) *For any étendue $\mathcal{E}$ over $\mathcal{S}$, there exists an* atomic *localic groupoid $\mathbb{G}$ in $\mathcal{S}$ (that is, one whose domain and codomain maps are local homeomorphisms) such that $\mathcal{E} \simeq \mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$.*

(iii) *For any bounded $\mathcal{S}$-topos $\mathcal{E}$ having a K-finite pre-bound over $\mathcal{S}$, there exists a proper localic groupoid $\mathbb{G}$ in $\mathcal{S}$ such that $\mathcal{E} \simeq \mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$.*

Proof (i) is immediate from 5.2.7 and the fact (5.1.1) that locally connected surjections are descent morphisms. (ii) similarly follows from the very definition of an étendue (5.2.4) and the fact that surjective local homeomorphisms are descent morphisms. (iii) follows from 5.2.9 and (the proper case of) 5.1.6. □

Example 5.2.12 The converse of 5.2.11(iii) is false. Let G be any profinite group which is not discrete: for example, the profinite completion of the additive group $\mathbb{Z}$ (that is, the inverse limit of its finite quotients in the category of topological groups). Then G is compact, and so $(G \rightrightarrows 1)$ is a proper localic groupoid. However, any K-finite object of $\mathbf{Cont}(G)$ has finite underlying set (by D5.4.12, since the forgetful functor $\mathbf{Cont}(G) \to \mathbf{Set}$ is an inverse image functor), and hence lies in the subcategory $\mathbf{Unif}(G)$ defined in A2.1.7; from this it is easy to see that it cannot be a bound, or even a pre-bound, for $\mathbf{Cont}(G)$ over $\mathbf{Set}$. We

remark also that the geometric morphism $\mathbf{Set} \to \mathbf{Cont}(G)$ is proper, by 5.3.6 below; thus $\mathbf{Cont}(G)$ also provides a counterexample to the converse of 5.2.9, as promised.

There is one further case of particular interest. If $\mathcal{E}$ is connected and atomic over $\mathcal{S}$, and it has a point $s\colon \mathcal{S} \to \mathcal{E}$, then the latter is automatically an open surjection by 3.5.6. In this case, we obtain a representation of $\mathcal{E}$ as $\mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$, where $\mathbb{G}$ is a localic groupoid in $\mathcal{S}$ satisfying $G_0 \cong 1$ – that is, a localic group. Thus we have

Theorem 5.2.13 *Any connected atomic $\mathcal{S}$-topos with a point is equivalent to one of the form $\mathbf{Cont}_{\mathcal{S}}(G)$, where G is an open localic group in $\mathcal{S}$ (i.e. one such that $G \to 1$ is open).* □

Remarks 5.2.14 (*a*) The assumption that the topos has a point cannot be omitted from 5.2.13: we give examples of (bounded) connected atomic morphisms with no sections in 5.4.12 below, and also in D3.4.14. (There are also unbounded examples, as we noted in 3.5.7.)

(*b*) The converse of 5.2.13 is also true: we already know this in the case when G is a topological group (i.e. its underlying locale is spatial), but it requires a little more proof if we merely assume G is an open locale. We shall provide the latter in 5.3.8 below.

(*c*) If the topos $\mathcal{E}$ in 5.2.13 is the classifying topos (over $\mathcal{S}$) for a geometric theory $\mathbb{T}$, then we may describe the group G appearing in the theorem a little more explicitly. The point s of course corresponds to a particular $\mathbb{T}$-model M in our base topos $\mathcal{S}$; thus, as a topos defined over $\mathcal{E}$, it may be thought of as classifying the theory of isomorphisms between the generic $\mathbb{T}$-model and our particular model M. Thus the pullback $\mathcal{S} \times_{\mathcal{E}} \mathcal{S}$ classifies the theory of a pair of such isomorphisms; but a pair of such ismorphisms must differ by some particular automorphism of M, and so we can identify this topos with $\mathbf{Sh}_{\mathcal{S}}(G)$, where G is the 'localic group of automorphisms' of M. At least in the case when $\mathcal{S} = \mathbf{Set}$ and the theory $\mathbb{T}$ is coherent, we may identify this with a topological group (that is, the locale G is spatial): it is the actual group of automorphisms of M, topologized with the 'topology of pointwise convergence' – i.e. the basic neighbourhoods of a given automorphism α are the sets

$$\{\beta \mid \beta \text{ agrees with } \alpha \text{ on } F\}$$

as F ranges over the finite subsets of M. We shall meet several particular examples of such groups in Section C5.4 below, and also in Section D3.4.

(*d*) A topos satisfying the hypotheses of 5.2.13 may have several non-isomorphic points, which may in turn have non-isomorphic automorphism groups: for example, the topos $\mathbf{Set}[\mathbb{D}_\infty]$ of D3.4.10 has one isomorphism class of points for each infinite cardinality in $\mathbf{Set}$, and these clearly have non-isomorphic automorphism groups. However, 5.2.13 tells us that all these groups are 'Morita-equivalent' in the sense that they give rise to equivalent toposes of continuous

actions. We shall explore this notion of Morita equivalence further in the next section.

Suggestions for further reading: Butz & Moerdijk [213, 214], Johnstone [524], Joyal & Tierney [560], Kock & Moerdijk [637, 638], Rosenthal [1047, 1048, 1051].

C5.3 Morita equivalence for groupoids

We have seen in the last section that every bounded $\mathcal{S}$-topos may be represented, up to equivalence, in the form $\mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$ for some localic groupoid $\mathbb{G}$ in $\mathcal{S}$. We have also seen, in B3.4.14(b), that every localic groupoid in $\mathcal{S}$ does give rise to a bounded $\mathcal{S}$-topos in this way. But there are clearly many different localic groupoids giving rise to any given $\mathcal{S}$-topos; so, for this representation to be useful, we need to have some control of the relation between localic groupoids which says that they give rise to equivalent categories of continuous actions. Borrowing a term from model theory (cf. D1.4.9), we shall call this relation *Morita equivalence.*

For most of this section, we shall be working with localic groups and groupoids in a fixed base topos $\mathcal{S}$, whose name we shall suppress whenever possible in order to simplify the notation; that is, we shall write $\mathbf{Cont}(\mathbb{G})$ and $\mathbf{Sh}(X)$ rather than $\mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$ and $\mathbf{Sh}_{\mathcal{S}}(X)$. However, apart from a few points where we refer explicitly to **Set**, all our arguments will be constructive, and so applicable in an arbitrary topos $\mathcal{S}$ – except that, as in the earlier sections of this chapter, our standing hypothesis that $\mathcal{S}$ has a natural number object will remain in force.

It will be recalled that we made some remarks about the problem of Morita equivalence, for the 'classical' case of a spatial group G, immediately following A2.1.6. Incidentally, from now on we shall tend to use the terms 'spatial group' and 'spatial groupoid', rather than 'topological group' and 'topological groupoid' to emphasize cases where the locales of objects and morphisms of a localic groupoid $\mathbb{G}$ are actually spaces (i.e. have enough points). Actually, there is a subtlety here: because the inclusion $\mathbf{Sob} \to \mathbf{Loc}$ does not preserve finite limits, a spatial group G (that is, an internal group in $\mathbf{Sp}$) may fail to be a localic group even if its underlying space is sober, because the multiplication defined on the product $G \times G$ in $\mathbf{Sp}$ may fail to extend to the locale product; and similarly for spatial groupoids. A counterexample is given by the additive group $\mathbb{Q}$ of rationals with the topology inherited from the Euclidean topology on $\mathbb{R}$, as we shall shortly see. However, we shall in practice feel free to ignore those spatial group(oid)s which are not localic group(oid)s – we shall not lose anything significant thereby.

We shall mainly concern ourselves with open localic groupoids (that is, those for which the domain and codomain maps $G_1 \rightrightarrows G_0$ are open); but many of our results have parallels for proper groupoids, and some of them apply to arbitrary groupoids in **Loc**. We begin our investigations with a result which strikingly

illustrates the difference between localic and spatial group(oid)s: every localic subgroup of a localic group (in **Set**) is closed. Of course, this immediately tells us that the spatial group $\mathbb{Q}$ cannot be a localic group, since it is a dense subgroup of the localic group $\mathbb{R}$ (and the latter *is* a localic group, since its local compactness ensures that the locale product $\mathbb{R} \times \mathbb{R}$ and the spatial product $(\mathbb{R} \times \mathbb{R})_p$ coincide; cf. 4.1.8). In fact we shall prove the result for localic groupoids, and then specialize to groups.

In general, given a locale Y over X, the notion of 'fibrewise closed sublocale of Y over X' obviously depends on the particular morphism $Y \to X$ which we have chosen, and not just on Y. However, if $\mathbb{G}$ is a localic groupoid, we shall tend to speak of sublocales of G_1 being fibrewise closed over G_0, without specifying explicitly whether our structure map $G_1 \to G_0$ is d_0 or d_1: the point is that, provided the sublocale in question is mapped to itself by the inverse map $i\colon G_1 \to G_1$ (as is clearly the case if it is the locale of morphisms of a subgroupoid of $\mathbb{G}$), then it doesn't matter which one we use, since i is an isomorphism $d_0 \to d_1$ in $\mathbf{Loc}/G_0$.

Theorem 5.3.1 *Let $\mathbb{G}$ be a localic groupoid, and let $\mathbb{H}$ be a localic subgroupoid of $\mathbb{G}$ (that is, we have a morphism $f\colon \mathbb{H} \to \mathbb{G}$ of localic groupoids such that both f_0 and f_1 are inclusions) such that the domain and codomain maps of $\mathbb{H}$ are open. Then $f_1\colon H_1 \rightarrowtail G_1$ is fibrewise closed over G_0.*

Proof We note first that we may reduce to the case when f_0 is an identity morphism, by replacing $\mathbb{G}$ by the groupoid $(\widehat{G_1} \rightrightarrows H_0)$ where

$$\begin{tikzcd}
\widehat{G_1} \arrow[r, tail] \arrow[d] & G_1 \arrow[d, "{(d_1, d_0)}"] \\
H_0 \times H_0 \arrow[r, tail, "f_0 \times f_0"] & G_0 \times G_0
\end{tikzcd}$$

is a pullback. For if $H_1 \rightarrowtail \widehat{G_1}$ is fibrewise closed over H_0, it is still fibrewise closed over G_0 since f_0 is an inclusion (cf. 1.2.14(i)); and the inclusion $\widehat{G_1} \rightarrowtail G_1$ is fibrewise closed over G_0, since it is the intersection of the pullbacks of $H_0 \rightarrowtail G_0$ along d_1 and d_0.

Now let $\overline{H_1}$ denote the fibrewise closure of H_1 in G_1 over G_0. We claim that $(\overline{H_1} \rightrightarrows G_0)$ can be made into a subgroupoid of $\mathbb{G}$; in particular, that the composition map $d_1\colon G_1 \times_{G_0} G_1 \to G_1$ restricts to a morphism $\overline{H_1} \times_{G_0} \overline{H_1} \to \overline{H_1}$

(the other conditions are easy). We note that we have pullbacks

$$\begin{tikzcd} H_1 \times_{G_0} \overline{H_1} \arrow[r, tail, "j \times 1"] \arrow[d] & \overline{H_1} \times_{G_0} \overline{H_1} \arrow[r, "d_2"] \arrow[d, "d_0"] & H_1 \arrow[d, "d_0"] \\ H_1 \arrow[r, tail, "j"] & \overline{H_1} \arrow[r, "d_1"] & G_0 \end{tikzcd}$$

where the domain and codomain maps $\overline{H_1} \rightrightarrows G_0$ are open by 3.1.14(ii), and the inclusion $j\colon H_1 \rightarrowtail \overline{H_1}$ is fibrewise dense over d_1; so, again by 3.1.14(ii), $j \times 1$ is fibrewise dense over $d_2\colon \overline{H_1} \times_{G_0} \overline{H_1} \to \overline{H_1}$, and hence over G_0 via either of the maps $d_1 d_2$ and $d_0 d_2$. But by applying 3.1.14(i) to the left-hand square above, we may also deduce that $j \times 1$ is fibrewise dense over G_0 via the third projection $d_0 d_0$. Similarly, the inclusion $\overline{H_1} \times_{G_0} H_1 \rightarrowtail \overline{H_1} \times_{G_0} \overline{H_1}$ is fibrewise dense over G_0, via any of the three projections; so their intersection $H_1 \times_{G_0} H_1 \rightarrowtail \overline{H_0} \times_{G_0} \overline{H_1}$ is also fibrewise dense over G_0, by 1.2.14(iii). But $d_1 : G_2 \to G_1$ may be regarded as a morphism over G_0, for either the first or the third projection $G_2 \to G_0$, and it restricts to a morphism $H_1 \times_{G_0} H_1 \to H_1$; so, by the functoriality of fibrewise closure (1.2.14(vi)), it also restricts to a morphism $\overline{H_1} \times_{G_0} \overline{H_1} \to \overline{H_1}$, as required.

Now that we know $\overline{H_1}$ defines a subgroupoid of $\mathbb{G}$, it follows that $\overline{H_2} = \overline{H_1} \times_{G_0} \overline{H_1}$ is invariant under the map $(id_0, d_1)\colon G_2 \to G_2$ (where $i\colon G_1 \to G_1$ is the inverse map of $\mathbb{G}$), and so is the sublocale $H_1 \times_{G_0} \overline{H_1}$. But the diagram

$$\begin{tikzcd} G_2 \arrow[r, "{(id_0, d_1)}"] \arrow[dr, "d_2"'] & G_2 \arrow[d, "d_1"] \\ & G_1 \end{tikzcd}$$

commutes; so, from the fact that $H_1 \times_{G_0} \overline{H_1} \rightarrowtail \overline{H_2}$ is fibrewise dense over d_2, we deduce that it is also fibrewise dense over d_1. Similarly, $\overline{H_1} \times_{G_0} H_1 \rightarrowtail \overline{H_2}$ is fibrewise dense over d_1, and hence so is their intersection $H_2 \rightarrowtail \overline{H_2}$. But $d_1\colon \overline{H_2} \to \overline{H_1}$ is (split) epic, so its restriction to H_2 is also epic, by 1.2.14(ii). And this restriction factors through $H_1 \rightarrowtail \overline{H_1}$, by assumption; so the latter must be an isomorphism, i.e. H_1 is fibrewise closed in G_1 over G_0. □

Specializing to the case of groups, we immediately obtain

Corollary 5.3.2 *Let G be a localic group. Then any localic subgroup of G which is open as a locale is weakly closed in G. In particular, if G is a localic group in a Boolean topos, then every localic subgroup of G is closed.* □

Remark 5.3.3 In particular, we note that for any localic group G the identity element $e: 1 \to G$ is a weakly closed sublocale of G; hence by 1.2.14(v) the diagonal $G \rightarrowtail G \times G$ is weakly closed, i.e. G is 'weakly Hausdorff'. However, we cannot strengthen this to 'Hausdorff' unless our underlying topos is Boolean. For a counterexample, let U be any non-complemented subterminal object in $\mathcal{S}$, and let G be the discrete locale on the pushout $1 \amalg_U 1$ in $\mathcal{S}$, equipped with a group structure as a quotient of the cyclic group of order two.

As another application of 5.3.1, we may deduce

Corollary 5.3.4 *If $R \rightrightarrows X$ is an open equivalence relation in* **Loc** *such that $R \rightarrowtail X \times X$ is an inclusion (and not merely monic), then it is effective.*

Proof We may regard R as a subgroupoid of the trivial groupoid $(X \times X \rightrightarrows X)$. Theorem 5.3.1 then says that $R \rightarrowtail X \times X$ is fibrewise closed over X; so the result follows from 5.1.10. □

Remark 5.3.5 The reader may be concerned about the apparent contradiction between 5.3.4 and the example of a non-effective equivalence relation in **LH**, using irrational rotations of the circle S^1, which we gave in A1.3.7. For this pair $R \rightrightarrows S^1$, the induced map $R \to S^1 \times S^1$ is merely monic in **Sp**, and not an inclusion; but if we replace R with its image $\overline{R} \rightarrowtail S^1 \times S^1$ (that is, if we re-topologize R as a subspace of $S^1 \times S^1$), we obtain an equivalence relation in **Sp** such that the projections $\overline{R} \rightrightarrows S^1$ are still open, but $\overline{R} \rightarrowtail S^1 \times S^1$ is fibrewise dense over S^1. Moreover, although $\overline{R} \rightrightarrows S^1$ is effective in **Sp**, it is not effective in **Loc** (or in **Sob**): its coequalizer in **Sp** is (indiscrete, and hence) not sober.

The answer to the conundrum is that, as a relation in **Loc**, $\overline{R} \rightarrowtail S^1 \times S^1$ is not transitive (unlike $R \to S^1 \times S^1$!), for exactly the same reason that $\mathbb{Q}$ fails to be a localic group: the localic pullback $\overline{R} \times_{S^1} \overline{R}$ is larger than the spatial pullback, and its image under the 'composition' map is the whole of $S^1 \times S^1$.

We recall that, for an arbitrary localic groupoid $\mathbb{G}$, the category $\mathbf{Cont}(\mathbb{G})$ of continuous actions of $\mathbb{G}$ is defined to be the full subcategory of the internal diagram category $[\mathbb{G}, \mathbf{Loc}]$ whose objects are those internal diagrams $((A \to G_0), \alpha)$ for which $A \to G_0$ is a local homeomorphism (i.e. an object of $\mathbf{Sh}(G_0)$). We defined this category, and saw that it is a topos, in B3.4.14(*b*); also, there is a surjective geometric morphism $d: \mathbf{Sh}(G_0) \to \mathbf{Cont}(\mathbb{G})$ whose inverse image is the forgetful functor. We note next that various properties of this morphism are inherited from the corresponding properties of the domain and codomain maps of $\mathbb{G}$.

Lemma 5.3.6 *Let $\mathbb{G}$ be a localic groupoid such that the domain and codomain maps $G_1 \rightrightarrows G_0$ are open (resp. locally connected, connected, atomic, proper, tidy). Then the same is true of the morphism d:* $\mathbf{Sh}(G_0) \to \mathbf{Cont}(\mathbb{G})$.

Proof First suppose d_0 and d_1 are open. To show that d is open, we shall show that its inverse image d^* commutes with universal quantification: let

$f\colon (A,\alpha) \to (B,\beta)$ be a morphism of $\mathbf{Cont}(\mathbb{G})$, and (A',α') a subobject of (A,α). It suffices to show that the universal quantification $B' = \forall_f(A') \rightarrowtail B$, computed in $\mathbf{Sh}(G_0)$, can be given the structure of a subobject of (B,β) in $\mathbf{Cont}(\mathbb{G})$, i.e. that the isomorphism $\beta\colon d_1^*(B) \to d_0^*(B)$ restricts to an isomorphism $d_1^*(B') \to d_0^*(B')$. But this is obvious from the fact that d_1^* and d_0^* commute with universal quantification; for we have $d_i^*(B') \cong \forall_{d_i^*(f)} d_i^*(A')$, and the isomorphism α identifies $d_1^*(A')$ and $d_0^*(A')$ as subobjects of $d_1^*(A) \cong d_0^*(A)$.

The arguments for the locally connected case are very similar, using the fact that a morphism is locally connected iff its inverse image commutes with Π-functors (3.3.1). For the connected case, let (A,α) and (B,β) be two objects of $\mathbf{Cont}(\mathbb{G})$; we have to show that every morphism $f\colon A \to B$ in $\mathbf{Sh}(G_0)$ is actually a morphism $(A,\alpha) \to (B,\beta)$, i.e. that it commutes with α and β. But the composite

$$d_0^*(A) \xrightarrow{\alpha^{-1}} d_1^*(A) \xrightarrow{d_1^*(f)} d_1^*(B) \xrightarrow{\beta} d_0^*(B)$$

must equal $d_0^*(g)$ for a unique $g\colon A \to B$; and since s_0^* sends both α and β (modulo coherence isomorphisms) to identity maps, we deduce on applying it to this composite that $f = g$. For the atomic case, we argue as in the locally connected case to show that d^* commutes with the formation of power objects.

In the proper case, we have to work rather harder. Let U be a subterminal object of $\mathbf{Sh}(G_0)$, and consider the object $\check{U} = d_{1*}d_0^*(U)$. Clearly, we have $\check{U} \leq d_{1*}s_{0*}s_0^*d_0^*(U) \cong U$; but we also have

$$d_1^*d_{1*}d_0^*U \cong d_{2*}d_1^*d_0^*(U) \cong d_{2*}d_0^*d_0^*(U) \cong d_0^*d_{1*}d_0^*(U) ,$$

using the weak Beck–Chevalley conditions for the pullback squares

$$\begin{array}{ccc} G_2 & \xrightarrow{d_2} & G_1 \\ \downarrow{\scriptstyle d_1} & & \downarrow{\scriptstyle d_1} \\ G_1 & \xrightarrow{d_1} & G_0 \end{array} \quad \text{and} \quad \begin{array}{ccc} G_2 & \xrightarrow{d_2} & G_1 \\ \downarrow{\scriptstyle d_0} & & \downarrow{\scriptstyle d_0} \\ G_1 & \xrightarrow{d_1} & G_0 \end{array}$$

and the identity $d_0d_0 = d_0d_1 : G_2 \to G_0$. So $\check{U}$, equipped with this isomorphism, may be thought of as an object of $\mathbf{Cont}(\mathbb{G})$ (since it is subterminal, the coherence conditions on the isomorphism are automatic). And any subterminal object V satisfying $V \leq U$ and $d_1^*(V) \cong d_0^*(V)$ also satisfies $V \leq d_{1*}d_1^*(V) \cong d_{1*}d_0^*(V) \leq d_{1*}d_0^*(U)$; so we may identify $U \mapsto \check{U}$ with the effect of d_* on subterminal objects. Now if (A,α) is any object of $\mathbf{Cont}(\mathbb{G})$, then on writing $(E_0 \to G_0)$ for the local homeomorphism corresponding to A, and $(E_1 \to G_1)$ for its pullback along either d_0 or d_1, we see that $(E_1 \rightrightarrows E_0)$ can be given the structure of a groupoid $\mathbb{E}$ in

Loc, and that $\mathbf{Cont}(\mathbb{G})/(A,\alpha)$ may be identified with $\mathbf{Cont}(\mathbb{E})$. So we may compute the effect of the direct image of $d/(A,\alpha)\colon \mathbf{Sh}(G_0)/A \to \mathbf{Cont}(\mathbb{G})/(A,\alpha)$ on subterminal objects in exactly the same way. Taking (A,α) to be the object of objects of a weakly directed internal poset in $\mathbf{Cont}(\mathbb{G})$ (so that A is the object of objects of a weakly directed internal poset in $\mathbf{Sh}(G_0)$), and using the fact that d_0^* and d_{1*} both preserve directed $\mathbf{Sh}(G_0)$-indexed unions of subobjects, we deduce that d_* preserves directed $\mathbf{Cont}(\mathbb{G})$-indexed unions of subobjects.

The argument in the tidy case is very similar, except that we work with arbitrary objects of the toposes concerned rather than subterminal objects. We leave the details to the reader. □

Let $\mathbb{G}$ be a localic groupoid in a topos $\mathcal{S}$. The functor $\mathfrak{BTop}/\mathcal{S} \to \mathfrak{Loc}(\mathcal{S})$ which sends a topos over $\mathcal{S}$ to its localic reflection is a left 2-adjoint (cf. 1.4.8), so it preserves colimits; hence, if we compute the (pseudo-)colimit of the simplicial locale $G_\bullet$ in $\mathfrak{Loc}(\mathcal{S})$, we obtain the localic reflection of $\mathbf{Cont}(\mathbb{G}) \to \mathcal{S}$. But, since $\mathfrak{Loc}(\mathcal{S})$ is a locally ordered 2-category (in which, moreover, all 2-isomorphisms are identities), there is no difference between the 'descent locale' $\mathbf{Desc}(G_\bullet)$ and the ordinary colimit of $G_\bullet$ computed in the 1-dimensional category $\mathbf{Loc}(\mathcal{S})$, i.e. the coequalizer of $G_1 \rightrightarrows G_0$ (which we shall denote by $\pi_0(\mathbb{G})$, and call the *locale of algebraic components* or *orbit locale* of $\mathbb{G}$). Thus we have

Lemma 5.3.7 *For any localic groupoid $\mathbb{G}$, there is a hyperconnected morphism* $\mathbf{Cont}(\mathbb{G}) \to \mathbf{Sh}(\pi_0(\mathbb{G}))$. *In particular, $\mathbb{G}$ is* algebraically connected *(that is, $G_1 \rightrightarrows G_0 \to 1$ is a coequalizer in* $\mathbf{Loc}(\mathcal{S})$*) iff* $\mathbf{Cont}(\mathbb{G})$ *is hyperconnected over $\mathcal{S}$.* □

We use the term 'algebraically connected' to distinguish this notion from the (topological) connectedness of $\mathbb{G}$ (i.e. of the domain and codomain maps $G_1 \rightrightarrows G_0$).

Using 5.3.6 and 5.3.7, we may now prove the converse of 5.2.13:

Corollary 5.3.8 *Let G be an open localic group in $\mathcal{S}$. Then* $\mathbf{Cont}(G)$ *is connected and atomic over $\mathcal{S}$ (and has a point).*

Proof By 5.3.6, the morphism $d\colon \mathcal{S} \simeq \mathbf{Sh}(G_0) \to \mathbf{Cont}(G)$ is an open surjection, so the fact that $\mathbf{Cont}(G) \to \mathcal{S}$ is atomic follows from 3.5.1(ii). The fact that it is connected (in fact hyperconnected) follows from 5.3.7. □

Remark 5.3.9 However, we could also prove 5.3.8 without appealing to 5.3.6 or 5.3.7, by constructing a connected atomic site of definition for $\mathbf{Cont}(G)$. We did this for a spatial group (in **Set**) in 2.2.4(c) (cf. also 3.5.9(b)); as we remarked there, the construction still works for a localic group, even though there may be no points in the cosets of its open subgroups. To see this, observe that if G is an open localic group, then for any open subgroup H of G the equivalence relation $G \times H \rightrightarrows G$ (given by the first projection and the multiplication) is effective by

5.3.4, from which it follows that the square

$$\begin{tikzcd} G \times H \arrow[r, "q\pi_1"] \arrow[d, "(\pi_1, m)"'] & G : H \arrow[d, "\Delta"] \\ G \times G \arrow[r, "q \times q"] & G : H \times G : H \end{tikzcd}$$

is a pullback (where q is the quotient map), and hence by 3.2.23 that $G : H$ (is open and) has open diagonal, i.e. it is discrete by 3.1.15. (So the left cosets of H exist as points of the discrete locale $G : H$, even though an individual coset may not contain any points of G. Also, if K is a particular left coset of H, then it is straightforward to verify that the image of $K \times K$ under the composite

$$G \times G \xrightarrow{1 \times i} G \times G \xrightarrow{m} G$$

is a subgroup isomorphic to H; thus we may speak of 'the conjugate gHg^{-1} corresponding to a left coset gH' even if no such g exists.)

$G : H$ also clearly inherits a left G-action from G itself, so it defines an object of $\mathbf{Cont}(G)$. Moreover, the objects of this type form an $\mathcal{S}$-indexed separating family for $\mathbf{Cont}(G)$: if E is any discrete locale equipped with a left action of G, then forming the equalizer of the projection $G \times E \to E$ and the action map yields an open sublocale of $G \times E$ (open because it is a pullback of the diagonal $E \to E \times E$) which we may think of as an E-indexed family of open subgroups of G, and which is such that the corresponding E-indexed family of quotients maps epimorphically to E. So we obtain a site of definition for $\mathbf{Cont}(G)$ by taking the internal full subcategory on the objects $G : H$; but it is easily seen that any morphism between two such objects is an epimorphism in $\mathbf{Cont}(G)$, and hence that every inhabited sieve is covering. So we have constructed a connected atomic site for $\mathbf{Cont}(G)$ over $\mathcal{S}$, in the sense of 3.5.8 (it is connected because it contains the terminal object $1 = G : G$ of $\mathbf{Cont}(G)$). Finally, it is easy to see that $\mathbf{Cont}(G)$ has a point, whose inverse image is the forgetful functor $\mathbf{Cont}(G) \to \mathcal{S}$.

Remark 5.3.10 We may also use the ideas of this section to give an alternative proof (at least for toposes with natural number objects) of the result of 2.4.14 that a bounded morphism whose diagonal is an inclusion must be localic. Let $p \colon \mathcal{E} \to \mathcal{S}$ be such a morphism, and let $f : \mathcal{F} \to \mathcal{E}$ be an open surjection with $\mathcal{F}$ localic over $\mathcal{S}$. Then $\mathcal{F} \times_{\mathcal{E}} \mathcal{F} \to \mathcal{F} \times_{\mathcal{S}} \mathcal{F}$ is a pullback of $\mathcal{E} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$, so it too is an inclusion. In terms of the localic groupoid $\mathbb{G}$ derived from this surjection, this says that $(d_1, d_0) \colon G_1 \to G_0 \times G_0$ is an inclusion in $\mathbf{Loc}(\mathcal{S})$; in particular, $G_1 \rightrightarrows G_0$ is an equivalence relation in $\mathbf{Loc}(\mathcal{S})$, and by 5.3.4 it is effective. Thus, if we form the coequalizer $q \colon G_0 \to \pi_0(\mathbb{G})$ of this pair in $\mathbf{Loc}(\mathcal{S})$, then q is an open surjection by the argument in the proof of 5.1.4, and hence $\mathbf{Sh}(G_0) \to \mathbf{Sh}(\pi_0(\mathbb{G}))$

is a descent morphism in $\mathfrak{Top}$ by 5.1.6. So we have equivalences $\mathbf{Sh}(\pi_0(\mathbb{G})) \simeq \mathbf{Desc}(\mathbf{Sh}(G_\bullet)) \simeq \mathcal{E}$.

We note next that the construction $\mathbf{Cont}(-)$ is indeed functorial on the category $\mathbf{Gpd}(\mathbf{Loc})$ of localic groupoids: if $f : \mathbb{G} \to \mathbb{H}$ is a functor between localic groupoids, then it extends to a morphism $f_\bullet : G_\bullet \to H_\bullet$ of simplicial locales in the obvious way, and hence induces a morphism of simplicial toposes $\mathbf{Sh}(G_\bullet) \to \mathbf{Sh}(H_\bullet)$. Since the operation of taking (pseudo-)colimits is functorial, this in turn induces a geometric morphism $\mathbf{Cont}(\mathbb{G}) \to \mathbf{Cont}(\mathbb{H})$, which we shall denote by $\mathbf{Cont}(f)$. It is not hard to verify that the inverse image of this morphism is simply the functor which sends a continuous $\mathbb{G}$-action (A, α) to $(f_0^*(A), f_1^*(\alpha))$ (modulo the appropriate coherence isomorphisms, whose names we suppress as usual). The direct image functor is not so easy to describe explicitly, so we shall avoid using it whenever possible.

$\mathbf{Cont}(-)$ is even 2-functorial. Here we have to be slightly careful, since we have two different sorts of 2-cells in our domain category $\mathbf{Gpd}(\mathbf{Loc})$, namely the 'topological' ones arising from the 2-dimensional structure of locales and the 'algebraic' ones arising from the 2-dimensional structure of groupoids. First, if we have two internal functors $f, g : \mathbb{G} \rightrightarrows \mathbb{H}$ such that $f_0 \leq g_0$ and $f_1 \leq g_1$ in the usual partial ordering on locale morphisms (cf. 1.1.6), then it is easy to see that $f_n \leq g_n$ for all n. Since these inequalities induce natural transformations $f_0^* \to g_0^*$ and $f_1^* \to g_1^*$, as we saw in 1.4.5, we obtain a natural transformation $\mathbf{Cont}(f)^* \to \mathbf{Cont}(g)^*$. But we also have the notion of internal natural transformation between internal functors $f, g : \mathbb{G} \rightrightarrows \mathbb{H}$ (which corresponds to the notion of simplicial homotopy, as we saw in B2.3.2). And these too induce geometric transformations: given $\theta : f \to g$, which we recall is a morphism $\theta_0 : G_0 \to H_1$ satisfying certain equations, and an object (A, α) of $\mathbf{Cont}(\mathbb{H})$, we see that $\theta_0^*(\alpha)$ is a morphism $f^*(A) \cong \theta_0^* d_1^*(A) \to \theta_0^* d_0^*(A) \cong g^*(A)$ in $\mathbf{Sh}(G_0)$, and straightforward verification shows that this is the component at (A, α) of a (necessarily invertible) natural transformation $\mathbf{Cont}(f)^* \to \mathbf{Cont}(g)^*$.

We may thus immediately deduce

Lemma 5.3.11 *Suppose we have an algebraic equivalence of groupoids between $\mathbb{G}$ and $\mathbb{H}$ – that is, a pair of functors $f : \mathbb{G} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{G}$ together with natural transformations from both fg and gf to the respective identity functors. Then the toposes* $\mathbf{Cont}(\mathbb{G})$ *and* $\mathbf{Cont}(\mathbb{H})$ *are equivalent.* □

But, as we emphasized back in Section A1.1, we have a weaker notion of equivalence of categories, which we there called 'weak equivalence': namely a functor which is full, faithful and essentially surjective on objects. What should this mean for a functor between localic groupoids? As usual, we call a functor

$f \colon \mathbb{G} \to \mathbb{H}$ *full and faithful* if

$$\begin{tikzcd} G_1 \arrow[r, "f_1"] \arrow[d, "(d_0, d_1)"] & H_1 \arrow[d, "(d_0, d_1)"] \\ G_0 \times G_0 \arrow[r, "f_0 \times f_0"] & H_0 \times H_0 \end{tikzcd}$$

is a pullback. For ordinary groupoids in $\mathcal{S}$, the statement that f is essentially surjective on objects would simply assert that the morphism $d_0\pi_1 \colon H_1 \times_{H_0} G_0 \to H_0$ is an epimorphism; but since we are working in **Loc**, where 'mere' epimorphisms are not even stable under pullback, we need to demand more than this. In fact it makes sense to demand that $d_0\pi_1$ should be a pullback-stable descent morphism in **Loc**. To achieve this, we shall in practice need to know that $d_0\pi_1$ is either an open surjection or a proper surjection; we shall call f an *open* (resp. *proper*) *weak equivalence* in the two cases.

Lemma 5.3.12 *If $f \colon \mathbb{G} \to \mathbb{H}$ is either an open or a proper weak equivalence of localic groupoids in $\mathcal{S}$, then the induced geometric morphism* **Cont**$(\mathbb{G}) \to$ **Cont**$(\mathbb{H})$ *is an equivalence.*

Proof Since $\mathbb{H}$ is a groupoid, the kernel-pair of $d_0 \colon H_1 \to H_0$ may be identified with $d_0, d_1 \colon H_2 \rightrightarrows H_1$; and, using the fact that f is full and faithful, an elementary diagram-chase shows that the kernel-pair of $d_0\pi_1 \colon H_1 \times_{H_0} G_0 \to H_0$ may be identified with

$$H_1 \times_{H_0} G_1 \underset{(d_1(1 \times f_1), d_1\pi_2)}{\overset{1 \times d_0}{\rightrightarrows}} H_1 \times_{H_0} G_0 \, .$$

Now if $(E, \alpha \colon d_1^* E \to d_0^* E)$ is an object of **Cont**$(\mathbb{G})$, then the object $\pi_2^*(E)$ of **Sh**$(H_1 \times_{H_0} G_0)$ may be equipped with descent data relative to this pair of maps by pulling back α along $\pi_2 \colon H_1 \times_{H_0} G_1 \to G_1$; so we obtain an object F of **Sh**(H_0) such that $(d_0\pi_1)^*(F) \cong \pi_2^*(E)$. Further, since the diagram

$$\begin{tikzcd} G_1 \arrow[r, "d_0"] \arrow[d, "(f_1, d_1)"'] & G_0 \arrow[d, "f_0"] \\ H_1 \times_{H_0} G_0 \arrow[r, "d_0\pi_1"] & H_0 \end{tikzcd}$$

commutes, we see that $d_0^* f_0^*(F) \cong d_1^*(E)$ in **Sh**(G_1); applying s_0^* to this, we get an isomorphism $f_0^*(F) \cong E$ in **Sh**(G_0).

To see that F can be equipped with an $\mathbb{H}$-action, we think of E as the domain of a local homeomorphism $E \to G_0$; then we can write $d_1^*(E)$ and $d_0^*(E)$ respectively as $G_1 \times_{G_0} E$ and $E \times_{G_0} G_1$, and we have a diagram in $\mathbf{Loc}(\mathcal{S})$

$$\begin{tikzcd}
H_1 \times_{H_0} G_1 \times_{G_0} E \arrow[r, shift left, "1 \times \overline{\alpha}"] \arrow[r, shift right, "(d_1(\pi_1, f_1\pi_2), \pi_3)"'] \arrow[d] & H_1 \times_{H_0} E \arrow[r] \arrow[d] & F \arrow[d] \\
H_1 \times_{H_0} G_1 \arrow[r, shift left, "1 \times d_0"] \arrow[r, shift right, "(d_1(1 \times f_1), d_1\pi_2)"'] & H_1 \times_{H_0} G_0 \arrow[r, "d_0\pi_1"] & H_0
\end{tikzcd}$$

where the vertical morphisms are local homeomorphisms and the top row is a pullback-stable coequalizer, by the proof of Proposition 5.1.4. (Here $\overline{\alpha}$ denotes the $\mathbb{G}$-action on E considered as a map $G_1 \times_{G_0} E \to E$, i.e. the composite of the isomorphism $G_1 \times_{G_0} E \to E \times_{G_0} G_1$ with the first projection.) Clearly, the first and second terms in the top row admit left $\mathbb{H}$-actions, and the morphisms between them are equivariant for this action; so it 'descends' to an action on F.

The construction above is easily seen to be functorial; so we have defined a functor $f_!\colon \mathbf{Cont}(\mathbb{G}) \to \mathbf{Cont}(\mathbb{H})$. It is also clear that the composite $f^*f_!$ is isomorphic to the identity. Finally, if $(E,\alpha) \cong f^*(F,\beta)$ for some object (F,β) of $\mathbf{Cont}(\mathbb{H})$, then we have pullback squares

$$\begin{tikzcd}
H_1 \times_{H_0} E \arrow[r] \arrow[d] & H_1 \times_{H_0} F \arrow[r, "\overline{\beta}"] \arrow[d] & F \arrow[d] \\
H_1 \times_{H_0} G_0 \arrow[r, "\pi_1"] & H_1 \arrow[r, "d_0"] & H_0
\end{tikzcd}$$

from which we may deduce that $f_!(E,\alpha)$ is isomorphic to (F,β). So the functors f^* and $f_!$ are inverse up to isomorphism. □

We may now prove a further representation theorem like those of Section C5.2, which is 'dual' to the representation of connected atomic toposes with a point given in 5.2.13. Recall that we say an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ is *Hausdorff* if p is a separated map, i.e. the diagonal $\mathcal{E} \to \mathcal{E} \times_{\mathcal{S}} \mathcal{E}$ is proper.

Corollary 5.3.13 *Let $\mathcal{E}$ be a hyperconnected Hausdorff $\mathcal{S}$-topos with a point. Then there exists a compact localic group G in $\mathcal{S}$ such that $\mathcal{E} \simeq \mathbf{Cont}_{\mathcal{S}}(G)$.*

We note that this is indeed dual to 5.2.13, since a geometric morphism is connected and atomic iff it is hyperconnected and has open diagonal, by 3.1.9(i), 3.5.4(i) and 3.5.14.

Proof Let $f: \mathcal{F} \to \mathcal{E}$ be an open surjection such that $\mathcal{F}$ is localic over $\mathcal{S}$; as in 5.2.6, we may choose this surjection so that the given point of $\mathcal{E}$ lifts to a point of $\mathcal{F}$. Applying 5.2.11, we obtain a representation of $\mathcal{E}$ as $\mathbf{Cont}_{\mathcal{S}}(\mathbb{H})$, where $\mathbb{H} = (H_1 \rightrightarrows H_0)$ is a localic groupoid which is open (since its domain and codomain maps are pullbacks of f), algebraically connected (by 5.3.7, since $\mathcal{E}$ is hyperconnected) and such that H_0 has a point $p: 1 \to H_0$. Moreover, the map $(d_0, d_1): H_1 \to H_0 \times H_0$ is proper, since it is a pullback of the diagonal of $\mathcal{E}$. Thus if we form the pullback of this morphism along $(p, p): 1 \to H_0 \times H_0$, we obtain a compact locale G, which clearly has a group structure inherited from the groupoid structure on $\mathbb{H}$. Also, we have a morphism of groupoids $\mathbb{G} = (G \rightrightarrows 1) \to \mathbb{H}$ defined by p and the inclusion $G \rightarrowtail H_1$, which is full and faithful by definition. And the equivalence relation on H_0 which is the image of $(d_0, d_1): H_1 \to H_0 \times H_0$ is closed in $H_0 \times H_0$, but has open projections since $\mathbb{H}$ is an open groupoid; so it is effective by 5.1.9, and hence must be the whole of $H_0 \times H_0$; in other words, $H_1 \to H_0 \times H_0$ is surjective. But $G_0 \times_{H_0} H_1 \to H_0$ is the pullback of this morphism along $p \times 1: H_0 \cong 1 \times H_0 \to H_0 \times H_0$; so it is also a proper surjection. Hence $\mathbb{G} \to \mathbb{H}$ is a proper weak equivalence of localic groupoids, and so by 5.3.12 we have $\mathbf{Cont}_{\mathcal{S}}(G) \simeq \mathbf{Cont}_{\mathcal{S}}(\mathbb{H}) \simeq \mathcal{E}$. □

Remarks 5.3.14 (*a*) Like 5.2.13, 5.3.13 has a converse. For if G is a compact localic group, then it follows from 5.3.6 that the canonical point $d: \mathcal{S} \simeq \mathbf{Sh}(G_0) \to \mathbf{Cont}(G)$ is proper; hence its kernel-pair $\mathcal{S} \times_{\mathbf{Cont}(G)} \mathcal{S}$ is a compact $\mathcal{S}$-topos, by the stability of proper maps under pullback. But we also have a pullback square

$$\begin{tikzcd}
\mathcal{S} \times_{\mathbf{Cont}(G)} \mathcal{S} \arrow[r] \arrow[d] & \mathbf{Cont}(G) \arrow[d, "\Delta"] \\
\mathcal{S} \arrow[r, "{(d,d)}"] & \mathbf{Cont}(G) \times_{\mathcal{S}} \mathbf{Cont}(G)
\end{tikzcd}$$

whose bottom edge is a proper surjection; so by 3.2.23 the diagonal Δ is proper, i.e. $\mathbf{Cont}(G)$ is a Hausdorff $\mathcal{S}$-topos. But we already know that it is hyperconnected and has a point, by 5.3.7. (Note also that the case of a discrete group G was covered in 3.2.24 above.)

(*b*) In the particular case when our base topos $\mathcal{S}$ is **Set**, we can say rather more about the group G which we obtain from 5.3.13. Since **Set** is Boolean, (the underlying locale of) any localic group is Hausdorff by 5.3.3. But any compact Hausdorff locale classically has enough points: this is a well-known result from locale theory (cf. [520, III 1.10]), but we may also obtain it from 4.1.13, which said that every compact Hausdorff locale is a retract of a coherent locale, plus D3.3.13 which says that every coherent topos has enough points. Thus we may take G to be a compact spatial group. Further, we noted after A2.1.6 that we may replace any spatial group G by its totally-disconnected reflection

G/G_e (where G_e is the connected component of the identity element) without changing the topos of continuous actions; and it is again a well-known result of general topology that a totally disconnected compact space is zero-dimensional (cf. [520, II 4.2]). Further, any zero-dimensional compact spatial group is prodiscrete, and indeed profinite, by [520, VI 2.9]. So we conclude that *the pointed hyperconnected Hausdorff toposes over* **Set** *are exactly the toposes of continuous actions of profinite groups.*

We may now deduce a result promised in Section C3.4:

Corollary 5.3.15 *Any proper separated map of toposes is tidy.*

Proof If f is proper and separated then so are both halves of its hyperconnected–localic factorization, by 3.2.16 and 3.2.25, so it suffices to prove the result separately under the hypothesis that f is either localic or hyperconnected. But the localic case was dealt with in 4.1.14. So suppose $f\colon \mathcal{F} \to \mathcal{E}$ is hyperconnected (hence proper) and separated. In general, it need not have a section; but its pullback $\pi_1\colon \mathcal{F} \times_{\mathcal{E}} \mathcal{F} \to \mathcal{F}$ has one, namely the diagonal, and it is still hyperconnected and separated. Hence by 5.3.13 we can represent it as $\mathbf{Cont}_{\mathcal{F}}(G)$ for some compact localic group G in $\mathcal{F}$, and it is therefore tidy by 3.4.1(b). But π_1 is the pullback of f along a proper surjection (namely, f itself); so f is tidy by 5.1.7. □

It is not only the weak equivalences considered in 5.3.12 that give rise to equivalences between toposes of continuous $\mathbb{G}$-actions. There is also a notion of 'completion' for localic groupoids, which has been christened *étale completion* by I. Moerdijk [831], which more or less by definition does not affect the topos of continuous $\mathbb{G}$-actions. To define it, let $\mathbb{G} = (G_1 \rightrightarrows G_0)$ be a localic groupoid, and let $d\colon \mathbf{Sh}(G_0) \to \mathbf{Cont}(\mathbb{G})$ be the geometric morphism whose inverse image is the forgetful functor. Then the square

$$\begin{array}{ccc} \mathbf{Sh}(G_1) & \xrightarrow{d_1} & \mathbf{Sh}(G_0) \\ \Big\downarrow{\scriptstyle d_0} & & \Big\downarrow{\scriptstyle d} \\ \mathbf{Sh}(G_0) & \xrightarrow{d} & \mathbf{Cont}(\mathbb{G}) \end{array}$$

commutes up to isomorphism, since by definition each object (A, α) of $\mathbf{Cont}(\mathbb{G})$ comes equipped with an isomorphism $\alpha\colon d_1^*(A) \to d_0^*(A)$ in $\mathbf{Sh}(G_1)$. We say that $\mathbb{G}$ is *étale-complete* if this square is a pseudo-pullback in $\mathfrak{BTop}/\mathcal{S}$; and we define the *étale completion* $\widehat{\mathbb{G}}$ of $\mathbb{G}$ to be $(\widehat{G_1} \rightrightarrows G_0)$, where $\mathbf{Sh}(\widehat{G_1}) \rightrightarrows \mathbf{Sh}(G_0)$ is the kernel-pair of d in $\mathfrak{BTop}/\mathcal{S}$. Note that this kernel-pair automatically has the structure of a groupoid; and that its domain is indeed localic over $\mathcal{S}$, since d is localic.

Lemma 5.3.16 *Let $\mathbb{G}$ be an open or proper localic groupoid. Then the canonical functor $\mathbb{G} \to \widehat{\mathbb{G}}$ induces an equivalence* **Cont**$(\mathbb{G}) \simeq$ **Cont**$(\widehat{\mathbb{G}})$. *Moreover, $\widehat{\mathbb{G}}$ is open (resp. locally connected, connected and locally connected, atomic, proper, tidy, tidy and connected) if $\mathbb{G}$ is.*

Proof Since d: **Sh**$(G_0) \to$ **Cont**$(\mathbb{G})$ is an open or proper surjection by 5.3.6, it is a descent morphism in $\mathfrak{BTop}/\mathcal{S}$; so **Cont**$(\mathbb{G})$ is equivalent to the category of objects of **Sh**(G_0) equipped with descent data relative to this morphism. But by definition the latter is precisely **Cont**$(\widehat{\mathbb{G}})$. The second assertion again follows from 5.3.6, plus the stability of the indicated classes of morphisms under pullback. □

Example 5.3.17 To gain some feeling for what étale completion means, we consider the particular case of a localic group G in **Set**. In this case, we saw in 5.3.9 that we have a site of definition for **Cont**(G) given by the category whose objects are the open subgroups H of G, with morphisms $H \to K$ labelled by left cosets gK of K such that $H \subseteq gKg^{-1}$, and equipped with the coverage in which every inhabited sieve covers. If it happens to be the case that every open subgroup of G contains an open normal subgroup, then we may use the Comparison Lemma to cut down this site to the full subcategory consisting of those open subgroups H which are normal, so that the corresponding object $G : H$ of **Cont**(G) admits a group structure. Of course, not every morphism in the site is a group homomorphism; but if we further cut down to those which are (i.e. those which are labelled by cosets containing the identity element of G), we obtain a codirected diagram of discrete groups and surjections, and if we take the inverse limit of this system in the category of localic groups, we obtain a group (the *pro-discrete completion* of G) whose category of discrete quotients is isomorphic to that of G (so that it gives rise to the same topos of continuous actions), and which may be shown to be the étale completion of G in this case.

In general, we cannot impose a group structure on the inverse limit of the discrete quotients $G : H$; the best we can get is a monoid structure. To define this, we may think of a 'point' of the inverse limit as a system of left cosets $(g_H H \mid H \in \mathcal{OS}(G))$ (where $\mathcal{OS}(G)$ denotes the poset of open subgroups of G), such that $g_H K = g_K K$ whenever $H \subseteq K$. Of course, this is only a manner of speaking, since the inverse limit may not have enough points; but what follows can be turned into a rigorous argument – cf. [831]. Then we define the product of two such systems $(g_H H \mid H \in \mathcal{OS}(G))$ and $(k_H H \mid H \in \mathcal{OS}(G))$ to be

$$(g_{k_H H k_H^{-1}} k_H H \mid H \in \mathcal{OS}(G)) .$$

It is readily verified that this is well-defined, and yields an associative multiplication on the inverse limit; however, not every 'point' of the limit has an inverse, in general. But if we cut down to the 'sublocale of invertible elements' of this monoid, we obtain a localic group which is the étale completion of G. We shall

not go into the details of the proof; the interested reader is referred to [831] and [833].

Finally in this section, we prove a theorem which says in effect that $\mathfrak{BTop}/\mathcal{S}$ may be regarded as a category of fractions of the full subcategory of $\mathbf{Gpd}(\mathbf{Loc}(\mathcal{S}))$ whose objects are étale-complete open localic groupoids in $\mathcal{S}$, obtained by inverting the open weak equivalences. In order to state this theorem formally, we should need to spend a good deal of time developing the theory of '2-categories of fractions', which seems hardly worth while given the amount of use we should make of it. So instead we shall merely state and prove the central result from which this representation may be deduced. (The reader who wishes to see a more precise formulation in terms of 2-categories of fractions is referred to [990].) We already know that every bounded $\mathcal{S}$-topos may be represented as $\mathbf{Cont}(\mathbb{G})$ for some étale-complete open localic groupoid $\mathbb{G}$, by 5.2.11(i) and 5.3.16; moreover, every open weak equivalence between such groupoids gives rise to an equivalence of toposes, by 5.3.12. It is not hard to verify that open weak equivalences (between open groupoids) are stable under composition and pullback, so that we should expect an arbitrary morphism $\mathbb{G} \to \mathbb{H}$ in the category of fractions to be representable by a span

$$\mathbb{G} \xleftarrow{\;g\;} \mathbb{K} \xrightarrow{\;h\;} \mathbb{H}\,,$$

where g is an open weak equivalence. In fact we shall prove that all geometric morphisms between toposes of continuous actions arise from spans of this type.

Theorem 5.3.18 *Let $\mathbb{G}$ and $\mathbb{H}$ be étale-complete open localic groupoids in $\mathcal{S}$, and let $f\colon \mathbf{Cont}(\mathbb{G}) \to \mathbf{Cont}(\mathbb{H})$ be a geometric morphism over $\mathcal{S}$. Then there exist an étale-complete open groupoid $\mathbb{K}$ and morphisms of groupoids g and h such that g is an open weak equivalence, and such that the diagram*

$$\begin{tikzcd} \mathbf{Cont}(\mathbb{K}) \arrow[d, "\mathbf{Cont}(g)"'] \arrow[dr, "\mathbf{Cont}(h)"] & \\ \mathbf{Cont}(\mathbb{G}) \arrow[r, "f"] & \mathbf{Cont}(\mathbb{H}) \end{tikzcd}$$

commutes up to isomorphism.

Proof Given f, we form the pullback

$$\begin{tikzcd} \mathcal{E} \arrow[rr, "h_0"] \arrow[d, "g_0"] & & \mathbf{Sh}(H_0) \arrow[d, "d"] \\ \mathbf{Sh}(G_0) \arrow[r, "d"] & \mathbf{Cont}(\mathbb{G}) \arrow[r, "f"] & \mathbf{Cont}(\mathbb{H}) \end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$. Since the right vertical morphism is localic, so is the left one; so we can write $\mathcal{E} \simeq \mathbf{Sh}(K_0)$ for some locale K_0. Moreover, $g_0 \colon \mathbf{Sh}(K_0) \to \mathbf{Sh}(G_0)$ is an open surjection, since $d \colon \mathbf{Sh}(H_0) \to \mathbf{Cont}(\mathbb{H})$ is an open surjection by 5.3.6. We define K_1 by the pullback diagram

$$\begin{tikzcd} K_1 \arrow[r, "g_1"] \arrow[d, "{(d_1, d_0)}"] & G_1 \arrow[d, "{(d_1, d_0)}"] \\ K_0 \times K_0 \arrow[r, "g_0 \times g_0"] & G_0 \times G_0 \end{tikzcd}$$

in $\mathbf{Loc}$; thus it is immediate that $\mathbb{K} = (K_1 \rightrightarrows K_0)$ inherits the structure of a groupoid from $\mathbb{G}$, and that $g \colon \mathbb{K} \to \mathbb{G}$ is full and faithful. Further, the domain and codomain maps of $\mathbb{K}$ are open, since they are pullbacks of those of $\mathbb{G}$, and $G_1 \times_{G_0} K_0 \to G_0$ is an open surjection, since we can factor it as the composite of $G_1 \to G_0$ and a pullback of $K_0 \to G_0$; so g is an open weak equivalence.

However, since $\mathbb{G}$ is an étale-complete groupoid, the geometric morphism $\mathbf{Sh}(G_1) \to \mathbf{Sh}(G_0 \times G_0) \simeq \mathbf{Sh}(G_0) \times_{\mathcal{S}} \mathbf{Sh}(G_0)$ is a pullback of the diagonal $\mathbf{Cont}(\mathbb{G}) \to \mathbf{Cont}(\mathbb{G}) \times_{\mathcal{S}} \mathbf{Cont}(\mathbb{G})$; and similarly for $\mathbb{H}$. So a straightforward diagram-chase shows that we could equivalently have defined K_1 by the pullback

$$\begin{tikzcd} \mathbf{Sh}(K_1) \arrow[rr, "h_1"] \arrow[d, "g_1"] & & \mathbf{Sh}(H_1) \arrow[d, "dd_1"] \\ \mathbf{Sh}(G_1) \arrow[r, "dd_1"] & \mathbf{Cont}(\mathbb{G}) \arrow[r, "f"] & \mathbf{Cont}(\mathbb{H}) \end{tikzcd}$$

in $\mathfrak{BTop}/\mathcal{S}$; and that the pair (h_0, h_1) defines a morphism of localic groupoids $\mathbb{K} \to \mathbb{H}$. The fact that $\mathbb{K}$ is étale-complete follows easily from the étale-completeness of $\mathbb{G}$ and the fact that g is full and faithful.

It thus remains only to verify that the diagram in the statement of the theorem commutes. Let (B, β) be an object of $\mathbf{Cont}(\mathbb{H})$, and write (A, α) for the object $f^*(B, \beta)$ of $\mathbf{Cont}(\mathbb{G})$. Then the (pseudo-)pullback square defining K_0 yields a canonical isomorphism $g_0^*(A) \cong h_0^*(B)$ in $\mathbf{Sh}(K_0)$; and, if we use $g_1^*(\alpha)$ and $h_1^*(\beta)$ to equip these two objects with $\mathbb{K}$-actions, it is easy to verify that the canonical isomorphism is in fact an isomorphism in $\mathbf{Cont}(\mathbb{K})$. So the result is established. □

We remark that 5.3.18 remains true, with exactly the same proof, if we substitute the word 'proper' for 'open' throughout. However, as we have already observed, not every bounded $\mathcal{S}$-topos is representable by a proper localic groupoid.

Suggestions for further reading: Isbell et al. [486], Johnstone [527, 528], Joyal & Moerdijk [554], Moerdijk [831–834, 837], Moerdijk & Vermeulen [858], Pronk [990].

C5.4 The Freyd representation

The present section would, in some respects, be more at home in Part D than in Part C. It is placed here because it leads to a representation theorem for (a large class of) Grothendieck toposes which can be presented in terms of localic groupoids; but the methods which lead up to this result are logical rather than geometric in character, and the references on which we shall rely are mostly to results proved in Part D. (Another difference between this section and the earlier ones in this chapter is that, for at least some of the results here, we shall need to assume that our base topos $\mathcal{S}$ is Boolean; in practice we shall take it to be the classical topos of sets, so that 'Grothendieck topos' is to be understood in its original and more restrictive sense.)

The representation theorem in question was first obtained by P. Freyd [374], and his motivation in proving it also came from logic: to obtain a better understanding of the way in which independence proofs in topos theory, of the sort which we discuss in Chapter F4, relate to their set-theoretic counterparts. Freyd observed that set-theoretic independence results are commonly proved using a combination of three techniques: passing from a given model of set theory to a 'forcing extension' corresponding to a complete Boolean algebra, passing to a 'Fraenkel–Mostowski model' corresponding to a topological group, and passing from a model of 'non-well-founded set theory' to its well-founded part, to recover the axiom of foundation. The first two of these, at least, have well-understood topos-theoretic analogues: they correspond to the passage from a topos $\mathcal{S}$ to the topos $\mathbf{Sh}_{\mathcal{S}}(B)$ of (canonical) sheaves on a complete Boolean algebra B in $\mathcal{S}$, and from $\mathcal{S}$ to the topos $\mathbf{Cont}_{\mathcal{S}}(G)$ of continuous G-sets for a topological group G in $\mathcal{S}$. In topos-theoretic terms, the third is a bit more nebulous, but it is at least an instance of the passage from a topos $\mathcal{E}$ to what Freyd called an *exponential variety*, that is a full subcategory closed under arbitrary products, coproducts, subobjects, quotients and power objects. (Such a subcategory is necessarily coreflective in $\mathcal{E}$, by an argument like that of A4.2.4(d); in fact, the exponential varieties in $\mathcal{E}$ are (up to equivalence) exactly the toposes $\mathcal{F}$ equipped with connected atomic geometric morphisms $\mathcal{E} \to \mathcal{F}$ – recall from 3.5.4(i) that a connected atomic morphism is hyperconnected.)

In considering topos-theoretic independence proofs, it obviously makes sense (since not all toposes are Boolean) to generalize the first of the three constructions above from complete Boolean algebras to complete Heyting algebras: that is, to allow the passage from $\mathcal{S}$ to any topos localic over $\mathcal{S}$ (cf. 1.4.7). But, at least in theory, we have a much larger generalization available: that from complete Heyting algebras to arbitrary sites. Do we in fact gain anything thereby? In

other words, do there exist Grothendieck toposes which cannot be obtained from **Set** by a combination of the three constructions described in the last paragraph? Freyd showed that the answer was negative: in fact (and rather surprisingly) we do not have to use all possible instances of the second construction, but only one particular one, corresponding to the group of all permutations of a countably infinite set. On the other hand, the first two constructions alone do not suffice to reach every Grothendieck topos from **Set** (though they do suffice to reach all Boolean Grothendieck toposes); the first part of this section will be devoted to characterizing the ones that can be reached. For this purpose, we need to recall some facts about decidable objects and their quotients.

We recall from A1.4.15 that an object A of a coherent category (in particular, of a Grothendieck topos) is said to be decidable if the diagonal subobject $\Delta_A : A \rightarrowtail A \times A$ has a complement. We recall also that the class of decidable objects in a Grothendieck topos is closed under finite products, arbitrary coproducts and arbitrary subobjects; and it follows easily from this that the class of quotients of decidable objects (i.e. those A for which there exists a decidable B and an epimorphism $B \twoheadrightarrow A$) is closed under finite products, arbitrary coproducts, arbitrary subobjects and arbitrary quotients. As we observed in A4.2.4(d), this implies that the full subcategory $\mathcal{E}_{qd}$ of quotients of decidable objects in $\mathcal{E}$ is a topos, and that there is a geometric morphism (in fact a hyperconnected one) $\mathcal{E} \to \mathcal{E}_{qd}$ whose inverse image is the inclusion. Here we shall be concerned (initially, at least) with the case when $\mathcal{E}_{qd}$ is the whole of $\mathcal{E}$.

Proposition 5.4.1 *For a Grothendieck topos $\mathcal{E}$, the following are equivalent:*

(i) *Every object of $\mathcal{E}$ is a quotient of a decidable object.*

(ii) *$\mathcal{E}$ has a separating set of decidable objects.*

(iii) *$\mathcal{E}$ has a bound over* **Set** *(in the sense of* B3.1.7*) which is decidable.*

(iv) *$\mathcal{E}$ has a pre-bound over* **Set** *(in the sense of* D3.2.3*) which is decidable.*

Proof (i) $\Rightarrow$ (iii): Given a separating set $\{G_i \mid i \in I\}$ for $\mathcal{E}$, let G be the coproduct of the G_i, and $F \twoheadrightarrow G$ an epimorphism with F decidable. Then G is a bound for $\mathcal{E}$, and hence so is F.

(iii) $\Rightarrow$ (iv) since any bound is a pre-bound; and (iv) $\Rightarrow$ (ii) because an object G is a pre-bound iff the subobjects of its finite powers form a separating set, and the latter are all decidable if G is.

Finally, (ii) $\Rightarrow$ (i) since every object can be expressed as a quotient of a coproduct of members of the separating set, and coproducts of decidable objects are decidable. □

If the equivalent conditions of 5.4.1 hold, we shall say that the topos $\mathcal{E}$ is *locally decidable.*

Examples 5.4.2 (a) Trivially, any Boolean topos is locally decidable.

(b) Any localic topos is locally decidable, since such a topos has 1 as a bound (or 0 as a pre-bound), and these are both decidable. More generally, if

$\mathcal{E}$ is locally decidable and $f: \mathcal{F} \to \mathcal{E}$ is a localic geometric morphism, then $\mathcal{F}$ is locally decidable, since f^* maps (pre-)bounds for $\mathcal{E}$ to (pre-)bounds for $\mathcal{F}$ by D3.2.4, and it preserves decidability.

(*c*) Let $\mathcal{C}$ be a small category. Then the functor category $[\mathcal{C}, \mathbf{Set}]$ is locally decidable iff every morphism of $\mathcal{C}$ is monic. For the latter condition is equivalent, by A1.4.16, to saying that the representable functors $\mathcal{C}(A, -)$, $A \in \text{ob}\, \mathcal{C}$, are all decidable, and they form a separating set for $[\mathcal{C}, \mathbf{Set}]$. Conversely, if $[\mathcal{C}, \mathbf{Set}]$ is locally decidable then the representable functors must be quotients of decidable objects; but, since they are projective, this forces them to be subobjects of decidable objects and hence decidable.

Remark 5.4.3 In connection with Example 5.4.2(*c*), we note that local decidability is 'dual' to the property of being an étendue, introduced in 5.2.4: for a small category $\mathcal{C}$, the functor category $[\mathcal{C}, \mathbf{Set}]$ is locally decidable iff $[\mathcal{C}^{\text{op}}, \mathbf{Set}]$ is an étendue. In particular, the classifying topos $\mathbf{Set}[\mathbb{D}] = [\mathbf{Set}_{fm}, \mathbf{Set}]$ for decidable objects (cf. D3.2.7) is locally decidable but not an étendue, and the classifying topos $\mathbf{Set}[\mathbb{K}] = [\mathbf{Set}_{fe}, \mathbf{Set}]$ for Kuratowski-finite objects (cf. D3.2.10) is an étendue but not locally decidable. (Further evidence that local decidability is dual to being an étendue may be found by comparing condition (v) of 5.4.4 below with Lemma 5.2.5.)

If $\mathcal{E}$ is locally decidable, then it follows from 5.4.1(iv) and D3.2.6 that there exists a localic geometric morphism from $\mathcal{E}$ to the classifying topos $\mathbf{Set}[\mathbb{D}]$ for decidable objects. The converse follows from 5.4.2(*b*) and 5.4.3. But in fact we can do better than this. If B is a (pre-)bound for $\mathcal{E}$, then so is $B \amalg N$, where N is the natural number object; and $B \amalg N$ is decidable if B is (recall that N is always decidable, by A2.5.11). But $B \amalg N$ is also infinite in the sense of D3.4.10, and so it can be classified by a (localic) geometric morphism from $\mathcal{E}$ to the classifying topos $\mathbf{Set}[\mathbb{D}_\infty]$ for infinite decidable objects. And in D3.4.10 we identified the latter topos as $\mathbf{Cont}(G_0)$, where G_0 is the group of all permutations of a countably infinite set (with its 'natural' topology). Thus we have established most of the following result:

Theorem 5.4.4 *For a Grothendieck topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is locally decidable.*

(ii) *There exists a localic geometric morphism from $\mathcal{E}$ to a Boolean topos.*

(iii) *There exists a localic geometric morphism $\mathcal{E} \to \mathbf{Cont}(G)$, for some spatial group G.*

(iv) *There exists a localic geometric morphism $\mathcal{E} \to \mathbf{Cont}(G_0)$, where G_0 is the (spatial) group of all permutations of $\mathbb{N}$.*

(v) *There exists a small site of definition $(\mathcal{C}, J)$ for $\mathcal{E}$ such that every morphism of $\mathcal{C}$ is epic.*

Proof (i) $\Rightarrow$ (iv) follows from the remarks above; (iv) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (ii) are trivial; and (ii) $\Rightarrow$ (i) follows from 5.4.2(a) and (b).

For (iii) $\Rightarrow$ (v), we recall that in 2.5.8(e) we constructed a site $(\mathcal{C}, J)$ for $\mathbf{Sh}_{\mathbf{Cont}(G)}(X)$, where X is an internal locale in $\mathbf{Cont}(G)$: the objects of $\mathcal{C}$ are pairs (H, U) where H is an open subgroup of G and U is an element of $\mathcal{O}(X)$ fixed by H, and morphisms $(H, U) \to (K, V)$ correspond to left cosets gK of K such that $H \subseteq gKg^{-1}$ and $U \subseteq gV$. A sieve R on (K, V) is covering iff the set of those $U \in \mathcal{O}(X)$ for which there exists $H \subseteq K$ such that $K: (H, U) \to (K, V) \in R$ has join V. It is easy to verify that every morphism of $\mathcal{C}$ is epic: for the composite of

$$(H, U) \xrightarrow{gK} (K, V) \xrightarrow{hL} (L, W)$$

is given by ghL, and clearly $ghL = gh'L$ implies $hL = h'L$.

Finally, (v) $\Rightarrow$ (i) follows from 5.4.2(b) and (c), since inclusions are localic. □

Corollary 5.4.5 *Any Boolean Grothendieck topos is equivalent to one of the form* $\mathbf{Sh}_{\mathbf{Cont}(G)}(X)$, *where* G *is a spatial group and* X *is a Boolean internal locale in* $\mathbf{Cont}(G)$ *(i.e. one such that* $\mathcal{O}(X)$ *is a complete Boolean algebra). Moreover, we may take* G *to be the particular group* G_0 *mentioned above.*

Proof A Boolean topos $\mathcal{E}$ is locally decidable, so by 5.4.4 it is representable as $\mathbf{Sh}_{\mathbf{Cont}(G_0)}(X)$ for some internal locale X in $\mathbf{Cont}(G_0)$. And the Booleanness of $\mathcal{E}$ forces X to be Boolean. □

The next remark provides a link with the ideas that we explored in the earlier sections of this chapter.

Remark 5.4.6 If G is a discrete group and X is a locale in $[G, \mathbf{Set}]$, then we may identify $\mathbf{Sh}_{[G,\mathbf{Set}]}(X)$ with the topos of continuous actions of the localic groupoid $(G \times X_0 \rightrightarrows X_0)$, where X_0 is the 'underlying locale in $\mathbf{Set}$' of X, i.e. the locale corresponding to the frame obtained by forgetting the G-action on $\mathcal{O}(X)$, and the domain and codomain maps $G \times X_0 \rightrightarrows X_0$ are respectively given by the second projection and the G-action. (Note that this is an atomic localic groupoid in the sense of 5.2.11(ii); that is, its domain and codomain maps are local homeomorphisms.)

We may similarly describe $\mathbf{Sh}_{\mathbf{Cont}(G)}(X)$, for an internal locale X in $\mathbf{Cont}(G)$, in terms of a localic groupoid, as follows. Since G acts on $\mathcal{O}(G)$ by translation, we may cut down to its 'continuous part'

$$\check{\mathcal{O}}(G) = \{U \in \mathcal{O}(G) \mid H \cdot U = U \text{ for some open subgroup } H\}$$

and regard the latter as an internal frame in $\mathbf{Cont}(G)$. If we denote the corresponding internal locale by $\check{G}$, then we have an action map $\check{G} \times X \to X$ in $\mathbf{Loc}(\mathbf{Cont}(G))$, and hence an internal localic groupoid $(\check{G} \times X \rightrightarrows X)$ in $\mathbf{Cont}(G)$. Taking the MacNeille completions of the corresponding frames, and

forgetting the G-actions, yields a localic groupoid in **Set**; and it may be shown that $\mathbf{Sh}_{\mathbf{Cont}(G)}(X)$ is equivalent to the topos of continuous actions of this localic groupoid.

If $\mathcal{E}$ is a Grothendieck topos which is not locally decidable, can we find a 'good approximation' to it which is locally decidable? In one direction, this is easy: we may now re-interpret A4.2.4(d) as follows.

Proposition 5.4.7 *For any Grothendieck topos $\mathcal{E}$, there is a locally decidable topos $\mathcal{E}_{qd}$ and a hyperconnected geometric morphism $\mathcal{E} \to \mathcal{E}_{qd}$ which is universal amongst geometric morphisms from $\mathcal{E}$ to locally decidable toposes.*

Proof The fact that $\mathcal{E}_{qd}$ is a topos, and that the inclusion $\mathcal{E}_{qd} \to \mathcal{E}$ is the inverse image of a geometric morphism, was proved in A4.2.4(d). This morphism is hyperconnected (and not merely connected) because $\mathcal{E}_{qd}$ is closed under arbitrary subobjects and quotients in $\mathcal{E}$. So it only remains to establish the universal property. But if $f \colon \mathcal{E} \to \mathcal{F}$ is a geometric morphism where $\mathcal{F}$ is locally decidable, then since f^* preserves decidability and epimorphisms it takes values in the subcategory $\mathcal{E}_{qd}$, and hence we have a factorization (unique up to unique isomorphism) of f through $\mathcal{E} \to \mathcal{E}_{qd}$. □

From 'the other side', the remark after 5.2.4 may be combined with 5.4.2(c) to yield a proof that every Grothendieck topos admits a hyperconnected morphism from a locally decidable one. But Freyd proved a stronger result: for every Grothendieck topos $\mathcal{E}$, there is a locally decidable $\mathcal{F}$ and a connected atomic morphism $\mathcal{F} \to \mathcal{E}$. Before proving this, however, we shall digress to consider the 'ordered' analogues of the results established so far in this section. For this purpose, we shall assume that our base topos **Set** satisfies the Linear Ordering Principle (the assertion that any set can be linearly ordered); this is well known to be strictly weaker than the full axiom of choice, but it implies the axiom of choice for families of finite sets.

We say an object A of a topos $\mathcal{E}$ is *orderable* if it can be given the structure of a model of the theory $\mathbb{L}$ of linearly ordered sets described in D3.4.11; and we say $\mathcal{E}$ is *locally orderable* if it has an orderable pre-bound. Clearly, an orderable object A is decidable, since the axioms of $\mathbb{L}$ imply that $A \times A$ is the disjoint union of the order-relation, its opposite and the diagonal; so any locally orderable topos is locally decidable. It is also clear that a coproduct $A_1 \amalg A_2$ of two orderable objects is orderable (with the ordering in which all of A_1 precedes all of A_2), and so is their product $A_1 \times A_2$ (with the lexicographic ordering). Hence, if B is an orderable pre-bound for $\mathcal{E}$, so is $(B \amalg 1) \times Q$, where Q is the object of rational numbers in $\mathcal{E}$ (cf. D4.7.1): but the latter pre-bound is actually a model of the theory $\mathbb{L}_\infty$ of dense linear orders without endpoints, considered in D3.4.11. Thus we obtain an analogue

of 5.4.4:

Theorem 5.4.8 *The following conditions on a Grothendieck topos $\mathcal{E}$ are equivalent:*

(i) *$\mathcal{E}$ is locally orderable.*

(ii) *Every object of $\mathcal{E}$ is a quotient of an orderable object.*

(iii) *There exists a localic geometric morphism from $\mathcal{E}$ to a topos in which every object is orderable.*

(iv) *There exists a localic geometric morphism from $\mathcal{E}$ to the topos $\mathbf{Cont}(G_1)$, where G_1 is the group of order-preserving permutations of $\mathbb{Q}$ (with its 'natural' topology).*

Proof (iii) $\Rightarrow$ (ii) follows from the fact that inverse image functors preserve orderability, and subobjects of orderable objects are orderable. (ii) $\Rightarrow$ (i) follows from the fact that any object mapping epimorphically to a pre-bound is a pre-bound. (i) $\Rightarrow$ (iv) follows from the remarks above, plus the fact that we may identify the classifying topos $\mathbf{Set}[\mathbb{L}_\infty]$ with $\mathbf{Cont}(G_1)$ (see D3.4.11). So, to complete the circle, it suffices to prove that every object of $\mathbf{Cont}(G_1)$ is orderable.

Since $\mathbf{Cont}(G_1)$ is atomic, we know that every object of it can be written as a set-indexed coproduct of atoms. Of course, in general the proof that an infinite coproduct of orderable objects is orderable requires the full axiom of choice in $\mathbf{Set}$, in order to choose a linear ordering of each summand (and a linear ordering of the set indexing the coproduct); however, it turns out that in $\mathbf{Cont}(G_1)$ each atom has a (nonzero) finite number of distinct linear orderings, so we need only the axiom of choice for families of finite sets – which, as we remarked above, follows from the Linear Ordering Principle.

By D3.4.11, we may also regard $\mathbf{Cont}(G_1)$ as the topos $\mathbf{Sh}(\mathbf{Ord}_{fm}^{\mathrm{op}}, J)$, where $\mathbf{Ord}_{fm}$ is the category of finite linearly ordered sets and order-preserving injections between them, and J is the atomic coverage on $\mathbf{Ord}_{fm}^{\mathrm{op}}$ (i.e. the coverage in which every inhabited sieve covers). In terms of this description, the atoms are the representable functors $A_n = \mathbf{Ord}_{fm}(n, -)$, where n denotes an n-element linearly ordered set. It is easy to see that A_n can be made into a subobject (in $n!$ possible ways) of $(A_1)^n$; and A_1 is the underlying object of the generic model of $\mathbb{L}_\infty$; hence every A_n is orderable. The fact that each A_n has finitely many linear orderings follows from the fact that $A_n \times A_n$ is a finite coproduct of atoms (since $(A_1)^{2n}$ is); so it has only finitely many subobjects. (In fact there are exactly $2^n n!$ linear orderings of A_n; for a proof of this, and further information about the structure of linearly ordered objects in $\mathbf{Cont}(G_1)$, see [529].) $\square$

Not every Boolean topos is locally orderable: for example, in $\mathbf{Cont}(G_0)$ the only orderable atom is the terminal object 1, from which it follows that the only orderable objects are the constant objects p^*I (where $p\colon \mathbf{Cont}(G_0) \to \mathbf{Set}$ is the

unique geometric morphism), and these do not include a pre-bound. (More generally, if G is any group containing non-identity elements of finite order, then $[G, \mathbf{Set}]$ is not locally orderable.) Perhaps more surprisingly, the conjunction of local orderability and Booleanness does not imply that every object is orderable; the following counterexample is essentially due to G. P. Monro [860] (see also [524]).

Example 5.4.9 Let Σ be the signature with one sort and three binary relation symbols $<$, $\equiv$ and $\#$ (which, as usual, we write in infix rather than prefix notation), and let $\mathbb{T}$ be the coherent theory over Σ with the following axioms:

(*a*) the axioms which say that $<$ is a linear order without endpoints;

(*b*) the axioms which say that $\equiv$ is an equivalence relation, and $\#$ is its complement;

(*c*) the sequents

$$\left(\top \vdash_{x_1,\ldots,x_n} (\exists y) \bigwedge_{i=1}^{n} (y \# x_i)\right)$$

for all $n \geq 0$, and

$$((x < y) \vdash_{x,y,z} (\exists w)((x < w) \wedge (w < y) \wedge (w \equiv z))) .$$

Thus a $\mathbb{T}$-model is a linearly ordered object without endpoints, equipped with a complemented equivalence relation which has infinitely many equivalence classes, and such that each equivalence class is dense in the linear order (so that, in particular, the linear order is dense). Thus $\mathbb{T}$ may be regarded as a propositional extension of $\mathbb{L}_\infty$, and so we have a localic morphism $\mathbf{Set}[\mathbb{T}] \to \mathbf{Set}[\mathbb{L}_\infty]$ by B4.2.12; in particular, $\mathbf{Set}[\mathbb{T}]$ is locally orderable. But it is easily seen that $\mathbb{T}$ satisfies the conditions of D3.4.6(ii), and so its classifying topos is Boolean; specifically, it is $\mathbf{Cont}(G_2)$ where G_2 is the automorphism group of the unique (up to isomorphism) countable $\mathbb{T}$-model M in $\mathbf{Set}$ (cf. 5.2.14(*c*)), and the generic model is M itself with the canonical action of G_2. Now consider the quotient $N = M/[\![\equiv]\!]_M$; of course, G_2 also acts continuously on this quotient, and (as an object of $\mathbf{Cont}(G_2)$) it is the quotient of the generic $\mathbb{T}$-model by $\equiv$. It is easy to see that, given any open subgroup H of G_2, we can find two elements x, y of N and an element of H which interchanges them; in fact we can take x and y to be any two $\equiv$-classes which do not contain elements fixed by H. It follows that N does not admit any linear ordering which is invariant under an open subgroup of G_2; hence, as an object of $\mathbf{Cont}(G_2)$, N is not orderable.

Remark 5.4.10 By analogy with 5.4.2(c), it is easy to see that a functor category $[\mathcal{C}, \mathbf{Set}]$ is locally orderable iff each representable functor $\mathcal{C}(U, -)$ admits a total ordering, iff each morphism of $\mathcal{C}$ is monic and we can totally order each hom-set of $\mathcal{C}$ in such a way that all mappings of the form $a \mapsto ba$ (for a fixed morphism b) are order-preserving. We note that this holds if $\mathcal{C}$ is the free category

on a directed graph $\mathcal{G}$, at least provided we have a total ordering of the arrows of $\mathcal{G}$, since we may then use reverse lexicographic ordering to order the morphisms of $\mathcal{C}$ (that is, finite formally-composable strings of arrows of $\mathcal{G}$) with a given domain. Thus in particular the étendue constructed in the proof of 5.2.3 is locally orderable; so that result may be strengthened to say that *every Grothendieck topos admits a hyperconnected morphism from a locally orderable étendue.*

We may also extend 5.4.8, by adding to its equivalent conditions one which corresponds to (v) of 5.4.4: a Grothendieck topos is locally orderable iff it is definable by a site $(\mathcal{C}, J)$ where $\mathcal{C}^{\mathrm{op}}$ satisfies the condition mentioned in the previous paragraph. For if our locally orderable topos $\mathcal{E}$ is represented as $\mathbf{Sh}_{\mathbf{Cont}(G_1)}(X)$, then we may use the Comparison Lemma 2.2.3 to cut down the site of 2.5.8(e) to its full subcategory on those objects (H, U) for which H is the stabilizer of a finite subset $\{q_1, \ldots, q_n\}$ of $\mathbb{Q}$. Now left cosets of such an H correspond to order-preserving injections $g \colon \{q_1, \ldots, q_n\} \to \mathbb{Q}$, and we may order these by imposing the lexicographic ordering on the n-tuples $(g(q_1), \ldots, g(q_n))$ (where for definiteness we suppose that the elements $q_1, \ldots, q_n$ are listed in increasing order). It is easily verified that this defines a total ordering on the hom-sets of $\mathcal{C}$ in such a way that composition with a fixed morphism is order-preserving.

We now return to the proof of Freyd's representation theorem.

Theorem 5.4.11 *Let $\mathcal{E}$ be a Grothendieck topos. Then, for $i = 0, 1$, there exist spans*

$$\mathcal{E} \xleftarrow{\;f_i\;} \mathcal{F}_i \xrightarrow{\;g_i\;} \mathbf{Cont}(G_i)$$

in $\mathfrak{BTop}/\mathbf{Set}$ such that f_i is connected atomic and g_i is localic, where G_0 (resp. G_1) is the group of all permutations of $\mathbb{N}$ (resp. the group of all order-preserving permutations of $\mathbb{Q}$), with its natural topology. Equivalently, $\mathcal{E}$ may be represented as an exponential variety in a locally decidable (resp. locally orderable) Grothendieck topos.

Proof We shall give two proofs of the theorem: the first is (a reworking of) Freyd's original proof, and the second is a more 'conceptual' proof, which is due to A. Joyal. (In fact the second proof is really no more than a translation of the first into the language of geometric theories: the relationship between the two is much like that between 5.2.1 and 5.2.8(c).) It is also worth noting that the case $i = 0$ of the theorem follows from the case $i = 1$, since the topos $\mathbf{Cont}(G_1)$ is Boolean and so admits a localic morphism to $\mathbf{Cont}(G_0)$ by 5.4.5; however, in both cases we shall give the proof for $i = 0$ in detail and merely sketch the (very similar) argument for $i = 1$.

(i) Let B be a pre-bound for $\mathcal{E}$ over $\mathbf{Set}$, that is, an object such that the subobjects of its finite powers form a generating set for $\mathcal{E}$. For technical reasons, we need to assume that B is well-supported, so that every injection $u \colon [m] \rightarrowtail [n]$ between finite sets induces an epimorphism $B^u \colon B^n \twoheadrightarrow B^m$ (here we are writing $[n]$ for the n-element set $\{0, 1, \ldots, n-1\}$). We shall take our site $(\mathcal{C}, J)$ for $\mathcal{E}$ to

be the full subcategory of $\mathcal{E}$ on the subobjects of finite powers of B, equipped with the coverage induced by the canonical coverage on $\mathcal{E}$.

We now define a new site $(\mathcal{D}, K)$ as follows: objects of $\mathcal{D}$ are diagrams

$$A \xrightarrow{f} A' \rightarrowtail^{g} B^n$$

in $\mathcal{E}$, where A and A' are objects of $\mathcal{C}$ and n is a natural number. We shall denote such an object by a quintuple (A, f, A', g, n). Morphisms $(A_1, f_1, A'_1, g_1, n_1) \to (A_2, f_2, A'_2, g_2, n_2)$ are pairs (h, u) where $h\colon A_1 \to A_2$ in $\mathcal{E}$ and $u\colon [n_2] \rightarrowtail [n_1]$ is an injection such that $B^u\colon B^{n_1} \to B^{n_2}$ maps the subobject g_1 into g_2 and the diagram

$$\begin{tikzcd} A_1 \arrow[r, "f_1"] \arrow[d, "h"] & A'_1 \arrow[r, tail, "g_1"] & B^{n_1} \arrow[d, "B^u"] \\ A_2 \arrow[r, "f_2"] & A'_2 \arrow[r, tail, "g_2"'] & B^{n_2} \end{tikzcd}$$

commutes. We have a functor $P\colon \mathcal{D} \to \mathcal{C}$ which sends (A, f, A', g, n) to A and (h, u) to h; it is easy to see that P is a fibration, since if we are given a morphism $h\colon C \to P(A, f, A', g, n)$ in $\mathcal{C}$ we have a prone lifting $(h, 1_{[n]})\colon (C, fh, A', g, n) \to (A, f, A', g, n)$. Also, P has a full and faithful right adjoint T given by $T(A) = (A \to 1 \rightarrowtail B^0)$; so if we define the coverage K on $\mathcal{D}$ by saying that P creates covers (i.e. that a sieve on an object is covering iff its image under P generates a J-covering sieve), then it follows from 3.5.10 that the fibration of sites (P, T) induces a connected atomic morphism $\mathbf{Sh}(\mathcal{D}, K) \to \mathbf{Sh}(\mathcal{C}, J) \simeq \mathcal{E}$.

Next, we observe that the full subcategory $\mathcal{D}'$ of $\mathcal{D}$ consisting of those objects (A, f, A', g, n) for which f is an identity morphism is K-dense in the sense of 2.2.1: for, given any object (A, f, A', g, n), we can choose a monomorphism $h\colon A \rightarrowtail G^m$ for some m, and then we have a morphism

$$\begin{tikzcd} A \arrow[r, "1"] \arrow[d, "1"] & A \arrow[r, tail, "{(h, gf)}"] & B^{m+n} \arrow[d, "B^u"] \\ A \arrow[r, "f"] & A' \arrow[r, tail, "g"] & B^n \end{tikzcd}$$

in $\mathcal{D}$ (where $u\colon [n] \to [m+n]$ is the map $j \mapsto m + j$), which clearly generates a K-covering sieve. Thus by 2.2.3 $\mathbf{Sh}(\mathcal{D}, K)$ is equivalent to $\mathbf{Sh}(\mathcal{D}', K')$, where K' is the coverage induced on $\mathcal{D}'$.

Now let L_n, for each n, denote the frame of subobjects of B^n in $\mathcal{E}$, and consider the sequence of embeddings

$$L_0 \rightarrowtail L_1 \rightarrowtail L_2 \rightarrowtail \cdots$$

induced by pullback along the projections $B^{n+1} \to B^n$ induced by the inclusions $[n] \rightarrowtail [n+1]$ in $\mathbf{Set}_f$. Let L be the colimit of this sequence (in $\mathbf{Set}$, or in the category of Heyting algebras). Given an element $g\colon A \rightarrowtail B^n$ of L_n for some n and a permutation u of $\mathbb{N}$, we may choose m such that u maps $[n]$ into $[m]$ and define $u \cdot f$ to be the subobject of B^m obtained by pulling back f along B^u; this clearly defines a continuous action of the group G_0 on L, in which L_n is exactly the subset of elements fixed by the subgroup $G_{0,n}$ of all permutations fixing $\{0, 1, \ldots, n-1\}$ pointwise. (The fact that L_n is exactly this subset, rather than merely contained in it, is a consequence of the fact that the morphisms B^u are epic in $\mathcal{E}$: if an object of $\mathrm{Sub}_{\mathcal{E}}(B^m)$, for some $m > n$, has the same pullback along B^u and B^v whenever u and v agree on $[n]$, then we may equip it with descent data relative to the epimorphism B^i, where $i\colon [n] \rightarrowtail [m]$ is the inclusion.)

In fact it is not hard to see that L is an internal frame in $\mathbf{Cont}(G_0)$, i.e. (cf. 2.5.8(e)) it is the 'continuous part' of its MacNeille completion, regarded as a frame in $[G_0, \mathbf{Set}]$. Moreover, if we construct the site described in 2.5.8(e) for the topos $\mathbf{Sh}_{\mathbf{Cont}(G_0)}(L)$ (but restricting its objects to those (H, U) for which H is one of the particular open subgroups $G_{0,n}$ – these clearly define a dense subcategory), we obtain a site equivalent to $(\mathcal{D}', K')$. Thus we have $\mathbf{Sh}(\mathcal{D}, K) \simeq \mathbf{Sh}(\mathcal{D}', K') \simeq \mathbf{Sh}_{\mathbf{Cont}(G_0)}(L)$, as required.

For the case $i = 1$, we proceed in exactly the same way, except that in the definition of the morphisms of $\mathcal{D}$ we restrict ourselves to pairs (h, u) for which u is an order-preserving injection, and we replace the sequential colimit in the definition of L by a directed colimit indexed by the poset of finite subsets of $\mathbb{Q}$.

(ii) For the second proof, we first note that, if $\mathcal{E}$ is a topos satisfying the conclusion of the theorem and $h\colon \mathcal{E}' \to \mathcal{E}$ is a localic geometric morphism, then $\mathcal{E}'$ also satisfies the conclusion; for we may form the pullbacks

$$\begin{tikzcd} \mathcal{F}'_i \arrow[r, "h_i"] \arrow[d, "f'_i"] & \mathcal{F}_i \arrow[d, "f_i"] \\ \mathcal{E}' \arrow[r, "h"] & \mathcal{E} \end{tikzcd}$$

in which f'_i is connected and atomic by 3.3.15 and 3.5.12, and h_i (and therefore the composite $g_i h_i$) is localic by B3.3.6. So, by D3.2.6, it suffices to prove the result for some particular topos $\mathbf{Set}[\mathbb{T}]$, where $\mathbb{T}$ is a single-sorted geometric theory such that every Grothendieck topos has a pre-bound which can be given

the structure of a $\mathbb{T}$-model. In fact, as in the second part of 5.2.8(*c*), we shall take $\mathbb{T}$ to be the theory $\mathbb{O}_1$ of inhabited objects.

We now proceed in a manner which is reminiscent of 5.4.9 above. For the case $i = 0$, we consider the theory $\mathbb{T}_0$ over the signature with two binary relation symbols $\#$ and $\equiv$, having axioms which say that $\#$ is the complement of the equality relation, that $\equiv$ is an equivalence relation, that the object under consideration is inhabited and, for each $n \geq 1$, the sequent

$$\left(\top \vdash_{x_1,\ldots,x_n} (\exists y)\left((y \equiv x_1) \wedge \bigwedge_{i=1}^{n} (y \# x_i)\right)\right)$$

which says that each $\equiv$-class is infinite. Let M denote the generic $\mathbb{T}_0$-model in $\mathbf{Set}[\mathbb{T}_0]$; then M is clearly decidable and infinite, so it is classified by a geometric morphism $\mathbf{Set}[\mathbb{T}_0] \to \mathbf{Set}[\mathbb{D}_\infty] \simeq \mathbf{Cont}(G_0)$, which is localic by D3.2.6. On the other hand, the quotient of M by the equivalence relation $[\![\equiv]\!]_M$ is clearly an inhabited object of $\mathbf{Set}[\mathbb{T}_0]$, so it is classified by a geometric morphism $\mathbf{Set}[\mathbb{T}_0] \to \mathbf{Set}[\mathbb{O}_1]$; we shall show that this morphism is connected and atomic.

Viewed as a topos over $\mathbf{Set}[\mathbb{O}_1]$, $\mathbf{Set}[\mathbb{T}_0]$ may be regarded as the classifying topos for the theory of 'G-indexed families of infinite decidable objects' where G is a particular object of the base topos, namely the generic inhabited object itself. (In fact, the pullback of $\mathbf{Set}[\mathbb{T}_0] \to \mathbf{Set}[\mathbb{O}_1]$ along an arbitrary localic morphism $f: \mathcal{E} \to \mathbf{Set}[\mathbb{O}_1]$ may similarly be regarded as classifying the theory of B-indexed families of decidable objects, where B is the pre-bound for $\mathcal{E}$ classified by f; so the argument which follows could be applied directly to this pullback. We note also that $\mathbf{Set}[\mathbb{T}_0] \to \mathbf{Set}[\mathbb{O}_1]$ is a (covariant) partial product in the sense of B4.4.14, namely the partial product of the opfibration $\mathbf{Set}[\mathbb{O}_1]/G \to \mathbf{Set}[\mathbb{O}_1]$ and the topos $\mathbf{Set}[\mathbb{D}_\infty]$; but that is not relevant to our present concerns.) It is easy to see, generalizing the arguments of D3.2.7 and D3.4.10, that a site for this topos may be obtained by taking the internal category $\mathbb{C}$ whose objects (indexed by the list object LG in $\mathbf{Set}[\mathbb{O}_1]$) are all morphisms $[p] \to G$ from finite cardinals to G, and whose morphisms are monomorphisms (over G) between their domains, and imposing the coverage T on $\mathbb{C}^{\mathrm{op}}$ in which every inhabited sieve covers. Now it is also easy to see that $(\mathbb{C}^{\mathrm{op}}, T)$ is a connected atomic site in the sense of 3.5.8; so $\mathbf{Set}[\mathbb{T}_0] \to \mathbf{Set}[\mathbb{O}_1]$ is connected and atomic.

For the case $i = 1$, we proceed similarly, using the theory $\mathbb{T}_1$ whose signature has two binary relation symbols $<$ and $\equiv$, and axioms which say that $<$ is a linear order without endpoints, that $\equiv$ is an equivalence relation, that the object under discussion is inhabited, and that each $\equiv$-class is dense in the linear order (cf. 5.4.9). Once again, the generic model M of $\mathbb{T}_1$ is classified by a localic geometric morphism $\mathbf{Set}[\mathbb{T}_1] \to \mathbf{Set}[\mathbb{L}_\infty] \simeq \mathbf{Cont}(G_1)$, and the quotient $M/[\![\equiv]\!]_M$ is classified by a connected atomic morphism $\mathbf{Set}[\mathbb{T}_1] \to \mathbf{Set}[\mathbb{O}_1]$. We omit the details, which are similar to those already given for $i = 0$. □

Remark 5.4.12 If $\mathcal{E}$ is any Grothendieck topos which is not locally decidable, then the covering maps $f_i: \mathcal{F}_i \to \mathcal{E}$ constructed in 5.4.11 provide examples of

(bounded) connected atomic morphisms with no sections (cf. 3.5.7 and 5.2.14(a)). We can see this directly, by observing that the geometric theories (relative to $\mathcal{E}$) occurring in the second proof of 5.4.11 cannot have models in $\mathcal{E}$; but it is also clear that a section of f_i, if it existed, would necessarily be localic (since its direct image functor would be faithful, cf. A4.6.2(a)), and so it would force $\mathcal{E}$ to be locally decidable. In the same way, taking $i = 1$ and $\mathcal{E} = \mathbf{Cont}(G_0)$, we get an example of such a morphism whose codomain is Boolean. In D3.4.14 we provide a (more complicated) example whose codomain is **Set**.

Suggestions for further reading: Freyd [374], Johnstone [521, 524, 529].

PART

TOPOSES AS '

)

HEORIES

D1

FIRST-ORDER CATEGORICAL LOGIC

D1.1 First-order languages

In this chapter we shall examine the relationship between topos theory and first-order predicate logic. The languages underlying this kind of logic contain formal expressions called *terms*, which denote individuals, and formal expressions called *formulae*, which denote predicates that make assertions about the individuals. Compound terms and formulae are built up using various logical operators. The expressions of such a language can be assigned meanings in a category: the terms are interpreted as morphisms and the formulae as subobjects, in a way that respects the logical structure of compound expressions. This categorical semantics requires various kinds of categorical structure to be present in the category, mirroring the logical operations which are allowed in the language, and will be described in the next section. In this section we shall describe the formal syntax of first-order languages.

Since our concern in this book is with toposes, we shall be considering those logical operations which have a meaningful interpretation using the categorical structure of toposes (both elementary and Grothendieck toposes): this is a very rich collection of logical operations. The formulae are built up using logical constants (true and false), connectives (conjunction, disjunction, implication and negation) and quantifiers (existential and universal), starting from some atomic relations (which will always include the distinguished relation of equality). The languages will be possibly *infinitary*, that is, allow infinite conjunctions or disjunctions; but they will be *first-order* rather than *higher-order*, meaning that quantification is over variable individuals rather than over subsets of individuals, functions of individuals, or iterates of such power and function constructions. (Higher-order predicate logic will be considered in Chapter D4.) However, we will allow individuals to be grouped into different *sorts*. Thus the languages considered in this chapter will contain expressions in *infinitary, first-order, many-sorted predicate logic with equality*. Each language may involve certain non-logical symbols denoting basic sorts and basic functions and relations of various kinds. This information constitutes the signature of the language.

Definition 1.1.1 A (first-order) *signature* Σ consists of the following data.

(*a*) A set Σ-Sort of *sorts*.

(*b*) A set Σ-Fun of *function symbols*, together with a map assigning to each $f \in \Sigma$-Fun its *type*, which consists of a finite non-empty list of sorts (with the last sort in the list enjoying a distinguished status): we write

$$f \colon A_1 \cdots A_n \to B$$

to indicate that f has type $A_1, \ldots, A_n, B$. (The number n is called the *arity* of f; in the case $n = 0$, f is more usually called a *constant* of sort B.)

(*c*) A set Σ-Rel of *relation symbols*, together with a map assigning to each $R \in \Sigma$-Rel its *type*, which consists of a finite list of sorts: we write

$$R \rightarrowtail A_1 \cdots A_n$$

to indicate that R has type $A_1, \ldots, A_n$. (The number n is called the *arity* of R; in the case $n = 0$, R is more usually called an *atomic proposition*.)

In some contexts, it is convenient to assume that the set Σ-Sort is equipped with finite products (that is, that we are given a set of 'ground sorts', and further sorts can be built up from these as (formal) finite products – including the empty product); this has the advantage that we can take all our function symbols and relation symbols to be unary. For most purposes, it is merely a matter of taste which approach one adopts in the first-order context; but when we come to higher-order logic in Chapter D4, it will be natural to adopt the second approach.

For each sort A of a signature Σ we assume given a supply of *variables* of sort A. (The variables of sort A are usually assumed to form a (fixed) countably infinite set V_A, but this doesn't really matter: the only important thing is that, in any context, we are free to introduce a new variable of sort A, i.e. one which hasn't already appeared in any of the terms or formulae under consideration.) Then we have

Definition 1.1.2 The collection of *terms* over Σ is defined recursively by the clauses below; simultaneously, we define the *sort* of each term and write $t \colon A$ to denote that t is a term of sort A.

(*a*) $x \colon A$, if x is a variable of sort A;

(*b*) $f(t_1, \ldots, t_n) \colon B$, if $f \colon A_1 \cdots A_n \to B$ is a function symbol and $t_1 : A_1, \ldots, t_n : A_n$.

If f is a constant (that is, a function symbol of arity 0), we commonly write f rather than $f(\,)$ for the term obtained by applying f (as in (*b*) above) to the empty string of terms.

Next, we turn to the *formulae*, that is the logical expressions denoting predicates; the following definition is complicated by the fact that we are simultaneously defining several different fragments of logic (that is, classes of formulae) with which we wish to work, but each individual clause in the definition is straightforward enough. (We should perhaps mention the long-standing tradition in logic whereby formulae are called 'well-formed formulae' or even 'wffs'; but since there is no (useful) notion of an 'ill-formed formula', this terminology seems redundant.)

Definition 1.1.3 Consider the following clauses for recursively defining a class F of *formulae* over Σ, together with, for each formula ϕ, the (finite) set $\mathrm{FV}(\phi)$ of *free variables* of ϕ.

(i) *Relations*: $R(t_1, \ldots, t_n)$ is in F, if $R \rightarrowtail A_1 \cdots A_n$ is a relation symbol and $t_1 \colon A_1, \ldots, t_n \colon A_n$ are terms; the free variables of this formula are all the variables occurring in some t_i. (Once again, if R has arity 0 we write simply R rather than $R(\,)$.)

(ii) *Equality*: $(s = t)$ is in F if s and t are terms of the same sort; $\mathrm{FV}(s = t)$ is the set of variables occurring in s or t (or both). (Sometimes it is convenient to distinguish between the equality predicates associated with different sorts; if we wish to do this, we shall write $(s =_A t)$, where A is the sort of s and t.)

(iii) *Truth*: $\top$ is in F; $\mathrm{FV}(\top) = \emptyset$.

(iv) *Binary conjunction*: $(\phi \wedge \psi)$ is in F, if ϕ and ψ are in F; $\mathrm{FV}(\phi \wedge \psi) = \mathrm{FV}(\phi) \cup \mathrm{FV}(\psi)$.

(v) *Falsity*: $\bot$ is in F; $\mathrm{FV}(\bot) = \emptyset$.

(vi) *Binary disjunction*: $(\phi \vee \psi)$ is in F, if ϕ and ψ are in F; $\mathrm{FV}(\phi \vee \psi) = \mathrm{FV}(\phi) \cup \mathrm{FV}(\psi)$.

(vii) *Implication*: $(\phi \Rightarrow \psi)$ is in F, if ϕ and ψ are in F; $\mathrm{FV}(\phi \Rightarrow \psi) = \mathrm{FV}(\phi) \cup \mathrm{FV}(\psi)$.

(viii) *Negation*: $\neg\phi$ is in F, if ϕ is in F; $\mathrm{FV}(\neg\phi) = \mathrm{FV}(\phi)$.

(ix) *Existential quantification*: $(\exists x)\phi$ is in F, if ϕ is in F and x is a variable; $\mathrm{FV}((\exists x)\phi) = \mathrm{FV}(\phi) \setminus \{x\}$. (Again, it is sometimes convenient to specify the sort of the variable being quantified; if we wish to do this, we write $(\exists x \colon A)\phi$.)

(x) *Universal quantification*: $(\forall x)\phi$ (or $(\forall x \colon A)\phi$, if it is necessary to specify the sort of x) is in F, if ϕ is in F and x is a variable; $\mathrm{FV}((\forall x)\phi) = \mathrm{FV}(\phi) \setminus \{x\}$.

(xi) *Infinitary disjunction*: $\bigvee_{i \in I} \phi_i$ is in F, if I is a set, ϕ_i is in F for each $i \in I$ and $\bigcup_{i \in I} \mathrm{FV}(\phi_i)$ is finite; in which case the latter set is $\mathrm{FV}(\bigvee_{i \in I} \phi_i)$.

(xii) *Infinitary conjunction*: $\bigwedge_{i \in I} \phi_i$ is in F, if I is a set, ϕ_i is in F for each $i \in I$ and $\bigcup_{i \in I} \mathrm{FV}(\phi_i)$ is finite; in which case the latter set is $\mathrm{FV}(\bigwedge_{i \in I} \phi_i)$.

Using these clauses, we make the following definitions.

(*a*) The set of *atomic formulae* over Σ is the smallest set closed under clauses (i) and (ii).

(*b*) The set of *Horn formulae* over Σ is the smallest set closed under clauses (i)–(iv).

(*c*) The set of *regular formulae* over Σ is the smallest set closed under clauses (i)–(iv) and (ix).

(*d*) The set of *coherent formulae* over Σ is the smallest set closed under clauses (i)–(vi) and (ix).

(*e*) The set of *first-order formulae* over Σ is the smallest set closed under clauses (i)–(x).

The final two definitions involve clauses (xi) and (xii), which are infinitary in nature and result in proper classes (rather than sets) of formulae.

(*f*) The class of *geometric formulae* over Σ is the smallest class closed under clauses (i)–(vi), (ix) and (xi).

(*g*) The class of *infinitary first-order formulae* over Σ is the smallest class closed under clauses (i)–(xii).

We recall that the names 'regular', 'coherent' and 'geometric' were all introduced in Chapter A1 as names for particular classes of categories with structure. The re-use of these names here is not accidental: as we shall see in the next section, regular (resp. coherent, geometric) formulae are precisely those which can be interpreted in regular (resp. coherent, geometric) categories. (Similarly, first-order formulae are those which can be interpreted in Heyting categories; we should perhaps have called the latter 'first-order categories', but that name is potentially confusing since every category is a model of a first-order theory – see 1.1.7(*e*) below.) There is one important class of formulae missing from the list above, namely *cartesian formulae*, which correspond in a similar way to cartesian categories: this class is in fact intermediate between those of Horn formulae and regular formulae (we allow certain instances of existential quantification, but not all of them), and before we can explain which existential quantifications are permitted we have to develop the notion of provability appropriate to cartesian logic. So the definition of this class is postponed until 1.3.4 below.

The variables x which occur in a formula ϕ, but which are within the scope of some quantifier $(\exists x)$ or $(\forall x)$, are called its *bound variables*. In practice we do not distinguish between two formulae if they are *α-equivalent*, that is, if they differ only in the names of bound variables. We note that the same variable may appear both bound and free in different parts of a compound formula; but for any ϕ there is an α-equivalent ϕ' in which this does not happen, and we normally suppose that the names of our bound variables have been chosen in such a way that these coincidences are avoided. A *closed formula* (also called a *sentence*) is one for which $\mathrm{FV}(\phi)$ is empty. A *closed term* is one which contains no variables.

It is quite possible, for a particular language, that there are sorts which possess no closed terms. Consequently, the use of variables of these sorts carries an implicit existential hypothesis. In classical set-based semantics, this implicit hypothesis is normally 'swept under the carpet' by imposing the requirement that the interpretation of any sort must be a *nonempty* set; but since we wish to interpret our languages in categories more general than **Set**, where we do not always have such a sharp distinction between 'emptiness' and 'possession of elements', we cannot afford to make any such general assumption. Therefore, we shall always keep a tally of which of these implicit hypotheses are present in any particular situation, via the notion of a 'context'.

Definition 1.1.4 A *context* is a finite list $\vec{x} = x_1, \ldots, x_n$ of distinct variables. The case $n = 0$ is allowed, being the empty context $[\,]$. If $\vec{x}$ is a context and y is a variable different from those occuring in $\vec{x}$, then $\vec{x}, y$ will denote the context obtained by appending y to the list $\vec{x}$. Similarly $\vec{x}, \vec{y}$ denotes the result of concatenating contexts $\vec{x}$ and $\vec{y}$ when they are disjoint. The *type* of a context $\vec{x}$ is the string of (not necessarily distinct!) sorts of the variables appearing in it.

We shall hardly ever consider terms and formulae by themselves, but rather in some context. We say a context $\vec{x}$ is *suitable* for a formula ϕ if all the free variables of ϕ occur in $\vec{x}$; a *formula-in-context* is an expression of the form $\vec{x}.\phi$, where ϕ is a formula and $\vec{x}$ is a suitable context for it. The *canonical context* for ϕ is the context consisting precisely of the distinct free variables of ϕ, listed in the order of their first appearance. Similarly, a *term-in-context* $\vec{x}.t$ is a term t together with a context $\vec{x}$ containing all the variables mentioned in t.

We shall frequently make use of the operation of substituting terms for variables in a formula. If $\vec{s} = s_1, \ldots, s_n$ is a list of (not necessarily distinct) terms, of the same length and type as the context $\vec{x}$, then

$$\phi[\vec{s}/\vec{x}]$$

will denote the formula (well-defined up to α-equivalence) resulting from simultaneously substituting s_i for each free occurrence of x_i in ϕ, for all $i \leq n$, after first changing the names of bound variables in ϕ if necessary to avoid capture of variables in $\vec{s}$ by any quantifiers in ϕ. (In practice, when making substitutions one normally takes care to ensure that the variables occurring in the terms to be substituted are disjoint from those which occur bound in the formula.) Similarly, $t[\vec{s}/\vec{x}]$ will denote the term resulting from simultaneously substituting each s_i for x_i in t.

The reader should be warned that the operation of simultaneous substitution, as defined above, is not the same as sequential substitution. In the formula $\phi[s_1/x_1][s_2/x_2]$, the term s_2 has been substituted for all occurrences of x_2 in s_1 (that is, in those copies of s_1 that have been substituted into ϕ), as well as the free occurrences of x_2 in ϕ; but in $\phi[s_1, s_2/x_1, x_2]$ it has not.

Next, we introduce the formal expressions which will serve as axioms for the logical theories we wish to consider.

Definition 1.1.5 By a *sequent* over a signature Σ we mean a formal expression of the form $(\phi \vdash_{\vec{x}} \psi)$, where ϕ and ψ are formulae over Σ and $\vec{x}$ is a context suitable for both of them. The intended interpretation of this expression is that ψ is a logical consequence of ϕ in the context $\vec{x}$, i.e. that any assignment of individual values to the variables in $\vec{x}$ which makes ϕ true will also make ψ true. We say a sequent $(\phi \vdash_{\vec{x}} \psi)$ is *regular* (*coherent*, ...) if both ϕ and ψ are regular (coherent, ...) formulae.

In full first-order logic, the general notion of sequent is not really needed, since the sequent $(\phi \vdash_{\vec{x}} \psi)$ expresses the same idea as $(\top \vdash_{[]} (\forall x_1) \dots (\forall x_n)(\phi \Rightarrow \psi))$, which we may simply identify with the formula on the right of the turnstile. However, the use of implication and universal quantification in the latter formula takes us out of any of the smaller fragments of logic that we wish to consider; so in these fragments it will be useful to have the notion of sequent available. In the opposite direction, it is sometimes helpful to consider sequents of the more general form $(\vec{\boldsymbol{\varphi}} \vdash_{\vec{x}} \psi)$, having a finite string $\vec{\boldsymbol{\varphi}} = (\phi_1, \dots, \phi_n)$ of hypotheses and a single conclusion; but since all the fragments of logic we wish to consider in this chapter will include finite conjunction, we shall have no real need of this extra generality.

Definition 1.1.6 By a *theory* over a signature Σ, we mean a set $\mathbb{T}$ of sequents over Σ, whose elements are called the (non-logical) *axioms* of $\mathbb{T}$. We say $\mathbb{T}$ is a *regular* (*coherent*, ...) *theory* if all the sequents in $\mathbb{T}$ are regular (coherent, ...).

In fact the above definition is a provisional one: we shall wish to identify two theories if they have the same logical consequences, and so what we have just defined is really the notion of a *presentation* or *axiomatization* of a theory. However, it is clearly impossible to give a more precise definition of what we mean by a theory until we have developed the notion of logical deduction in Section D1.3. The reason for introducing Definition 1.1.6 at this stage is to enable us to give the list of examples of theories with which we now conclude this section, so as to give the reader some idea of the expressive power of the various fragments of first-order logic that we have introduced.

Examples 1.1.7 (*a*) By an *algebraic theory*, we mean one whose signature Σ has a single sort and no relation symbols (apart from equality), and whose axioms are all of the form $(\top \vdash_{\vec{x}} \phi)$ where ϕ is an atomic formula $(s = t)$ and $\vec{x}$ is its canonical context. Such theories can be used to express the notion of a set (or, more generally, an object of a category) A equipped with 'algebraic structure' in the form of finitary operations $A^n \to A$, satisfying 'universal' equations: examples include monoids, groups, abelian groups, rings, modules (over a given ring), lattices, Heyting algebras (cf. A1.5.11), and so on.

(*b*) Slightly more generally, we may consider *many-sorted algebraic theories*, which are subject to the same restrictions as in (*a*) except that their signatures may have more than one sort. A well-known example is the theory of ring–module pairs: this has two sorts R and M, whose individuals are to be interpreted as

elements of the ring and of the module respectively, and (for example) scalar multiplication becomes a function symbol of type $R, M \to M$. Another example is the theory of directed graphs: this has two sorts V and A (for vertices and arrows), two function symbols of type $A \to V$ (for source and target), and no axioms. (More generally, for any small category $\mathcal{C}$ the theory of diagrams of shape $\mathcal{C}$ can be expressed as a many-sorted algebraic theory.)

(*c*) A familiar example of a theory whose signature involves relation symbols rather than function symbols is the theory of partially ordered sets. The signature for this theory has one sort A, one relation symbol $\leq\, \rightarrowtail A, A$ (which we write, as usual, in infix rather than prefix notation: $(s \leq t)$ denotes the atomic formula $\leq (s, t)$), and no function symbols. The axioms are

$$\begin{gathered}(\top \vdash_x (x \leq x)),\\ (((x \leq y) \wedge (y \leq x)) \vdash_{x,y} (x = y)), \text{ and}\\ (((x \leq y) \wedge (y \leq z)) \vdash_{x,y,z} (x \leq z)).\end{gathered}$$

It is clear that this theory is a Horn theory. More generally, we can axiomatize ordered algebraic structures such as ordered groups and ordered rings, by taking a signature which includes both the relation symbol $\leq$ and the appropriate function symbols, and writing down the appropriate Horn sequents as axioms. On the other hand, the theory of totally ordered sets, which is obtained from that of partially ordered sets by adding the sequent

$$(\top \vdash_{x,y} ((x \leq y) \vee (y \leq x))),$$

is (coherent but) not a Horn theory.

(*d*) Again within the realm of Horn theories, we can axiomatize the theories of various classes of algebraic structures such as torsion-free abelian groups. For this particular example, we take the algebraic theory of abelian groups (written additively), and add the Horn sequents

$$((nx = 0) \vdash_x (x = 0))$$

for every integer $n > 1$, where nx of course denotes the term

$$((\cdots(x + x) + \cdots) + x)$$

in which the variable x appears n times.

(*e*) The theory of categories, unlike that of posets, is not a Horn theory (of course, we are not in a position to prove at present that it cannot be axiomatized by Horn sequents, but we shall see how to do so eventually – cf. 2.4.7 below), but it is a regular theory (and we shall see in Section D1.3 that it is actually a cartesian theory). It can be expressed over a signature with a single sort, but it is more convenient to take one with two sorts O and M (for objects and morphisms), three function symbols $\mathrm{id}\colon O \to M$, $\mathrm{dom}\colon M \to O$ and $\mathrm{cod}\colon M \to O$, and one relation symbol $T \rightarrowtail M, M, M$ (the intended meaning

of $T(x,y,z)$ being that z is the composite of x and y). The axioms involving T include

$$(T(x,y,z) \vdash_{x,y,z} ((\mathrm{dom}(x) = \mathrm{cod}(y)) \wedge (\mathrm{cod}(x) = \mathrm{cod}(z)) \wedge (\mathrm{dom}(y) = \mathrm{dom}(z))),$$
$$((T(x,y,z) \wedge T(x,y,w)) \vdash_{x,y,z,w} (z = w)), \text{ and}$$
$$((\mathrm{dom}(x) = \mathrm{cod}(y)) \vdash_{x,y} (\exists z)T(x,y,z));$$

we leave it to the reader to work out what the other axioms should say.

(*f*) Another example of a regular theory is the theory of divisible abelian groups: this is obtained from the theory of abelian groups by adding the axioms

$$(\top \vdash_x (\exists y)(ny = x))$$

for all $n > 1$. Unlike the theory of categories, this is not a cartesian theory (but if we add the torsion-freeness axioms from example (*d*), we do get a cartesian theory); again, the proof of this fact will come later.

(*g*) A local ring is commonly defined by ring-theorists as a nontrivial commutative ring (with 1) having a unique maximal ideal; this description appears to lie outside the realm of first-order logic since it contains an implicit quantification over ideals, but it is well known to be equivalent to the condition that the non-invertible elements form an additive subgroup (in fact an ideal), which is clearly first-order. If we 'turn round' this condition to get rid of the negations, it becomes a coherent theory: to the algebraic theory of commutative rings with 1, we add the axioms $((0 = 1) \vdash_{[]} \bot)$ (to ensure nontriviality), and

$$((\exists z)((x+y)z = 1) \vdash_{x,y} ((\exists z)(xz = 1) \vee (\exists z)(yz = 1))) .$$

(Here we have used the customary ring-theoretic notation for terms built up from variables using the function symbols for addition and multiplication.)

(*h*) Similarly, though a field is customarily described as a nontrivial ring in which every nonzero element is invertible (which would suggest the non-coherent sequent $(\neg(x = 0) \vdash_x (\exists y)(xy = 1))$), the theory of fields can be presented coherently: to the axioms for commutative rings with 1, plus the nontriviality axiom mentioned in (*g*), we add the coherent sequent

$$(\top \vdash_x ((x = 0) \vee (\exists y)(xy = 1))) .$$

The reader should be warned that this sequent, though classically equivalent to the non-coherent one mentioned earlier, is constructively stronger; and since we wish to interpret our logic in categories whose subobject lattices may not be Boolean, we shall have to employ constructive rather than classical logic. Thus we have two constructively inequivalent theories of fields, whose models in **Set** (or, more generally, in any Boolean coherent category) happen to coincide, but which 'look different' in a general Heyting category. (In fact, these two theories by no means exhaust the possible ways of defining a field; see [505].) However, it is a consequence of the completeness theorem for geometric theories, which

we shall prove in 3.1.16 below, that if two geometric theories (over the same signature) are classically equivalent (i.e. have the same models in all Boolean toposes), then they have the same models in all toposes.

(*i*) It is remarkable how few of the first-order theories encountered in the practice of mathematics fail to be (at least classically equivalent to) coherent theories. (There is a reason for this – see 1.5.13 below.) One that does so is the theory of posets in which each element lies below some maximal element: to the theory described in (*c*) above, we add the axiom

$$(\top \vdash_x (\exists y)((x \leq y) \wedge (\forall z)((y \leq z) \Rightarrow (y = z))))$$

which is clearly first-order but not coherent. (Once again, it is far from clear at present that this theory *cannot* be axiomatized by coherent sequents, but we shall eventually be able to prove that this is so – cf. 2.4.10.)

(*j*) As an example of a theory which is geometric but not (finitary) first-order, we mention the theory of torsion abelian groups: this is obtained from the algebraic theory of abelian groups by adding the sequent

$$\Big(\top \vdash_x \bigvee_{n>1} (nx = 0)\Big) .$$

In a similar vein, the theory of fields of a given finite characteristic p is first-order and even coherent (just add the sequent $\big(\top \vdash_{[]} (p1 = 0)\big)$ to the theory in (*h*)), but the theory of fields of finite characteristic is (geometric but) not first-order.

(*k*) Another instructive example of a geometric theory is the theory of finite sets. To axiomatize this, we take one sort A and one n-ary relation symbol R_n for each $n \geq 0$. Then, for each n, we take the axiom

$$\Big(R_n(x_1, \ldots x_n) \vdash_{x_1, \ldots, x_n, y} \bigvee_{i=1}^{n} (y = x_i)\Big)$$

which says that if an n-tuple of individuals satisfies the relation R_n then it exhausts the members of (the set interpreting) the sort. (The case $n = 0$ of this axiom is $(R_0 \vdash_y \perp)$, which says that if R_0 holds then the interpretation of the sort must be empty.) Then we add the axiom

$$\Big(\top \vdash_{[]} \bigvee_{n \geq 0} (\exists x_1) \cdots (\exists x_n) R_n(x_1, \ldots, x_n)\Big) .$$

Finally, to ensure that the interpretations of the R_n are uniquely determined by that of the sort (i.e. that R_n holds for *all* n-tuples which exhaust the elements of the sort, and not just for some of them), we must add the axioms

$$(R_n(x_1, \ldots, x_n) \vdash_{x_1, \ldots, x_n} R_m(x_{f(1)}, \ldots, x_{f(m)}))$$

whenever $f \colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ is a surjection, and

$$((R_n(x_1, \ldots, x_n) \wedge (x_i = x_j)) \vdash_{x_1, \ldots, x_n} R_{n-1}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n))$$

whenever $1 \leq i < j \leq n$.

(*l*) An example of an infinitary first-order theory which is not geometric is given by the theory of metric spaces. This may be axiomatized over a signature with one sort A and a family of binary relation symbols $R_\epsilon \rightarrowtail A, A$ indexed by positive real numbers ϵ; the intended interpretation of $R_\epsilon(x, y)$ is that the distance between x and y is (strictly) less than ϵ. Among the axioms, we have things like

$$(R_\epsilon(x,y) \vdash_{x,y} R_\delta(x,y)) \quad \text{whenever } \epsilon < \delta,$$
$$\Big(R_\epsilon(x,y) \vdash_{x,y} \bigvee_{\delta<\epsilon} R_\delta(x,y)\Big) \quad \text{for all } \epsilon, \text{ and}$$
$$((R_\epsilon(x,y) \wedge R_\delta(y,z)) \vdash_{x,y,z} R_{\epsilon+\delta}(x,z))$$

(the last of these expresses the triangle inequality), all of which are geometric; but the key axiom

$$\Big(\bigwedge_{\epsilon>0} R_\epsilon(x,y) \vdash_{x,y} (x=y)\Big),$$

which distinguishes metric spaces from pseudometric spaces, requires an infinitary conjunction.

(*m*) We conclude by briefly mentioning the class of *propositional theories.* A propositional signature is one which has no sorts; it follows that it cannot have any function symbols either, and the only relation symbols are atomic propositions (i.e. have arity 0). Of course, in a propositional signature we have no variables, and so we can forget about contexts – the only context is the empty one. Since there can be no quantifiers either, the distinction between Horn formulae/sequents/theories and regular formulae/sequents/theories disappears, and we call them *cartesian formulae/sequents/theories.* Propositional theories are useful for describing subsets of a given structure with particular properties. For example, given a meet-semilattice L, we have a propositional theory of filters in L, presented in a signature with one atomic proposition F_a (to be thought of as the assertion that a is in the filter) for each $a \in L$; the axioms are $(\top \vdash F_1)$ (where 1 is the top element of L), and all sequents of the form $(F_a \vdash F_b)$ with $a \leq b$ in L, or

$$((F_a \wedge F_b) \vdash F_{a\wedge b})$$

for arbitrary $a, b \in L$. This is clearly a cartesian theory; if L is a lattice, we may also obtain a coherent propositional theory of prime filters of L, by adding the axiom $(F_0 \vdash \bot)$ and all instances of

$$(F_{a\vee b} \vdash (F_a \vee F_b)) \,.$$

And if L is a complete lattice, we may axiomatize completely prime filters of L (which play an important rôle in the theory of locales – see C1.2.2) as a geometric

propositional theory, by adding the infinitary analogue of the last axiom-scheme:

$$\Big(F_a \vdash \bigvee_{i \in I} F_{a_i}\Big)$$

whenever a is the join of the family $(a_i \mid i \in I)$ in L. (These three examples are 'typical' in a very strong sense, as we shall see in 1.4.14 below.)

Suggestions for further reading: Kock & Reyes [640], Makkai & Reyes [790], Reyes [1004].

D1.2 Categorical semantics

This section describes the interpretation in categories of the expressions of a first-order language. This material lies at the heart of the relationship between category theory and predicate logic, and is a direct generalization of the traditional definition (due to A. Tarski) of satisfaction of first-order formulae in ordinary set-valued structures.

Definition 1.2.1 (*a*) Let $\mathcal{C}$ be a category with finite products and Σ a signature. A Σ-*structure* M in $\mathcal{C}$ is specified by the following data:

(i) A function assigning to each sort A in Σ-Sort, an object MA of $\mathcal{C}$. This function is extended to finite strings of sorts by defining $M(A_1, \ldots, A_n) = MA_1 \times \cdots \times MA_n$ (and setting $M([\,])$, where $[\,]$ denotes the empty string, equal to the terminal object 1 of $\mathcal{C}$).

(ii) A function assigning to each function symbol $f \colon A_1 \cdots A_n \to B$ in Σ-Fun a morphism $Mf \colon M(A_1, \ldots, A_n) \to MB$ in $\mathcal{C}$.

(iii) A function assigning to each relation symbol $R \rightarrowtail A_1 \cdots A_n$ in Σ-Rel a subobject $MR \rightarrowtail M(A_1, \ldots, A_n)$ in $\mathcal{C}$.

(*b*) The Σ-structures in $\mathcal{C}$ are the objects of a category Σ-$\mathbf{Str}(\mathcal{C})$ whose morphisms are the Σ-structure *homomorphisms*. Such a homomorphism $h \colon M \to N$ is specified by a collection of morphisms $h_A \colon MA \to NA$ in $\mathcal{C}$ indexed by the sorts of Σ and satisfying the following two conditions:

(iv) For each function symbol $f \colon A_1 \cdots A_n \to B$ in Σ-Fun, the diagram

$$\begin{tikzcd}
M(A_1, \ldots, A_n) \arrow[r, "Mf"] \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"] & MB \arrow[d, "h_B"] \\
N(A_1, \ldots, A_n) \arrow[r, "Nf"] & NB
\end{tikzcd}$$

commutes.

(v) For each relation symbol $R \rightarrowtail A_1 \cdots A_n$ in Σ-Rel, there is a commutative diagram in $\mathcal{C}$ of the form

$$\begin{tikzcd} MR \arrow[r, tail] \arrow[d] & M(A_1, \ldots, A_n) \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"] \\ NR \arrow[r, tail] & N(A_1, \ldots, A_n) \end{tikzcd} .$$

Identities and composition in Σ-$\mathbf{Str}(\mathcal{C})$ are defined componentwise from those in $\mathcal{C}$.

We note in passing that if $\mathcal{D}$ is another category with finite products, then any functor $T: \mathcal{C} \to \mathcal{D}$ which preserves finite products and monomorphisms induces a functor Σ-$\mathbf{Str}(T): \Sigma$-$\mathbf{Str}(\mathcal{C}) \to \Sigma$-$\mathbf{Str}(\mathcal{D})$ in the obvious way; and any natural transformation $\alpha: T_1 \to T_2$ between two such functors induces a natural transformation Σ-$\mathbf{Str}(T_1) \to \Sigma$-$\mathbf{Str}(T_2)$. Thus the construction Σ-$\mathbf{Str}(-)$ is 2-functorial.

Remark 1.2.2 There is an alternative 'looser' way of interpreting first-order languages in a category, which is sometimes known as the *Brouwer–Heyting–Kolmogorov interpretation* (or BHK interpretation, for short). For the BHK interpretation, we interpret sorts and function symbols exactly as in 1.2.1, but a relation symbol $R \rightarrowtail A_1 \cdots A_n$ is interpreted as an equivalence class of (not necessarily monic) morphisms $MR \to M(A_1, \ldots, A_n)$, where two morphisms $f: X \to M(A_1, \ldots, A_n)$ and $g: Y \to M(A_1, \ldots, A_n)$ are considered equivalent if they become isomorphic in the preorder reflection of $\mathcal{C}/M(A_1, \ldots, A_n)$, i.e. if there exist morphisms $f \to g$ and $g \to f$ in $\mathcal{C}/M(A_1, \ldots, A_n)$. It is possible to develop all the ideas of this section for the BHK interpretation, in parallel with the development which we shall give for the 'Tarski interpretation'; but we don't need to do this, since it is easy to see using A1.3.10(d) that a 'BHK structure' for a signature Σ in a category $\mathcal{C}$ is really the same thing as a Tarski structure for Σ in the regularization $\mathbf{Reg}(\mathcal{C})$ of $\mathcal{C}$, in which the sorts of Σ happen to have been interpreted by objects in the image of the canonical embedding $I: \mathcal{C} \to \mathbf{Reg}(\mathcal{C})$. Therefore, we shall not pursue this interpretation further.

In a Σ-structure in a category with finite products, the function symbols are interpreted as morphisms of an appropriate kind in the category. This interpretation extends to the terms over Σ (or rather, to the terms-in-context) in a natural way.

Definition 1.2.3 Let M be a Σ-structure in a category $\mathcal{C}$ with finite products. If $\vec{x}.t$ is a term-in-context over Σ (with $\vec{x} = x_1, \ldots, x_n$, $x_i : A_i$ $(i = 1, \ldots, n)$ and

$t\colon B$, say), then a morphism

$$[\![\vec{x}\,.\,t]\!]_M\colon M(A_1,\ldots,A_n)\to MB$$

in $\mathcal{C}$ is defined recursively by the following clauses:

(*a*) If t is a variable, it is necessarily x_i for some unique $i\leq n$, and then $[\![\vec{x}\,.\,t]\!]_M=\pi_i$, the ith product projection.

(*b*) If t is $f(t_1,\ldots,t_m)$ (where $t_i\colon C_i$, say), then $[\![\vec{x}\,.\,t]\!]_M$ is the composite

$$M(A_1,\ldots,A_n)\xrightarrow{([\![\vec{x}\,.\,t_1]\!]_M,\ldots,[\![\vec{x}\,.\,t_m]\!]_M)}M(C_1,\ldots,C_m)\xrightarrow{Mf}MB\,.$$

When it is clear which particular Σ-structure is being referred to, we shall omit the subscript M from $[\![\vec{x}\,.\,t]\!]$.

Implicit in the above definition is the use of composition in the category to interpret the operation of substituting a term for a variable in a term. That is, we have the following *Substitution Property*:

Lemma 1.2.4 *Let $\vec{y}$ be a suitable context for a term $t\colon C$ (where $y_i\colon B_i$, say), and let $\vec{s}$ be a string of terms of the same length and type as $\vec{y}$. Further, let $\vec{x}$ be a suitable context for each s_i. Then $[\![\vec{x}\,.\,t[\vec{s}/\vec{y}]]\!]$ is the composite*

$$M(A_1,\ldots,A_n)\xrightarrow{([\![\vec{x}\,.\,s_1]\!],\ldots,[\![\vec{x}\,.\,s_m]\!])}M(B_1,\ldots,B_m)\xrightarrow{[\![\vec{y}\,.\,t]\!]}MC\,.$$

Proof Straightforward induction over the structure of t. □

We remark that 1.2.4 covers the case where we wish to make substitutions for only some of the variables in $\vec{y}$, since we can take the remaining s_i to be y_i. In particular, we may consider the case where $\vec{x}$ contains $\vec{y}$ and each s_i is y_i; then $([\![\vec{x}\,.\,s_1]\!],\ldots,[\![\vec{x}\,.\,s_m]\!])$ is just a product projection morphism, π, and we have the so-called *Weakening Property*: $[\![\vec{x}\,.\,t]\!]=[\![\vec{y}\,.\,t]\!]\circ\pi$ whenever $\vec{y}\subseteq\vec{x}$. We see from the Weakening Property that knowledge of $[\![\vec{y}\,.\,t]\!]$, where $\vec{y}$ is the canonical context for t, determines $[\![\vec{x}\,.\,t]\!]$ for any other context $\vec{x}$ suitable for t. Occasionally, we shall simplify our notation further by writing $[\![t]\!]$ for the interpretation of t in its canonical context.

Another important consequence of Definition 1.2.3 is that the 'naturality' with respect to the interpretations of function symbols, which was built into the definition of a homomorphism of Σ-structures, extends to the interpretations of terms-in-context:

Lemma 1.2.5 *Let $h\colon M\to N$ be a homomorphism of Σ-structures in a category $\mathcal{C}$ with finite products, and let $\vec{x}\,.\,t$ be a term-in-context over Σ (where $t\colon B$,*

$\vec{x} = x_1, \ldots, x_n$ *and* $x_i \colon A_i$, *say). Then the diagram*

$$\begin{tikzcd} M(A_1, \ldots, A_n) \arrow[r, "{[\![\vec{x}\,.\,t]\!]_M}"] \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"] & MB \arrow[d, "h_B"] \\ N(A_1, \ldots, A_n) \arrow[r, "{[\![\vec{x}\,.\,t]\!]_N}"] & NB \end{tikzcd}$$

commutes.

Proof Again, this is an easy induction over the structure of t. □

We turn next to the interpretation of formulae in a Σ-structure. Once again, we interpret formulae-in-context rather than abstract formulae; also, the range of logical operators we are able to interpret will depend on the amount of categorical structure possessed by the category $\mathcal{C}$ in which the structure lives.

Definition 1.2.6 Let M be a Σ-structure in a category $\mathcal{C}$ with (at least) finite limits. A formula-in-context $\vec{x}\,.\,\phi$ over Σ (where $\vec{x} = x_1, \ldots, x_n$ and $x_i \colon A_i$, say) will be interpreted as a subobject

$$[\![\vec{x}\,.\,\phi]\!]_M \rightarrowtail M(A_1, \ldots, A_n)$$

according to the following recursive clauses (where, as before, we drop the subscript M if the structure under consideration is obvious):

(i) If ϕ is $R(t_1, \ldots, t_m)$ where R is a relation symbol (of type $B_1, \ldots, B_m$, say), then $[\![\vec{x}\,.\,\phi]\!]$ is the pullback

$$\begin{tikzcd} {[\![\vec{x}\,.\,\phi]\!]} \arrow[r] \arrow[d, tail] & MR \arrow[d, tail] \\ M(A_1, \ldots, A_n) \arrow[r, "{([\![\vec{x}\,.\,t_1]\!], \ldots, [\![\vec{x}\,.\,t_m]\!])}"] & M(B_1, \ldots, B_m) \end{tikzcd} .$$

(ii) If ϕ is $(s = t)$, where s and t are terms of sort B, then $[\![\vec{x}\,.\,\phi]\!]$ is the equalizer of

$$\begin{tikzcd} M(A_1, \ldots, A_n) \arrow[r, shift left, "{[\![\vec{x}\,.\,s]\!]}"] \arrow[r, shift right, "{[\![\vec{x}\,.\,t]\!]}"'] & MB \end{tikzcd} .$$

(Equivalently, $[\![\vec{x}\,.\,\phi]\!]$ is the pullback of the diagonal $MB \rightarrowtail MB \times MB$ along $([\![\vec{x}\,.\,s]\!], [\![\vec{x}\,.\,t]\!])$. Thus we may think of equality between terms of sort B as a special relation symbol of type B, B, which differs from

other relation symbols in that its interpretation in a Σ-structure is not an arbitrary subobject, but is required to be the diagonal.)

(iii) If ϕ is $\top$, then $[\![\vec{x}\,.\,\phi]\!]$ is the top element of $\mathrm{Sub}(M(A_1,\ldots,A_n))$.

(iv) If ϕ is $(\psi \wedge \chi)$, then $[\![\vec{x}\,.\,\phi]\!]$ is the intersection (= pullback)

$$\begin{tikzcd} {[\![\vec{x}\,.\,\phi]\!]} \arrow[r, tail] \arrow[d, tail] & {[\![\vec{x}\,.\,\psi]\!]} \arrow[d, tail] \\ {[\![\vec{x}\,.\,\chi]\!]} \arrow[r, tail] & M(A_1,\ldots,A_n) \end{tikzcd} .$$

(v) If ϕ is $\bot$, and $\mathcal{C}$ is a coherent category, then $[\![\vec{x}\,.\,\phi]\!]$ is the bottom element of $\mathrm{Sub}(M(A_1,\ldots,A_n))$.

(vi) Similarly, if ϕ is $(\psi \vee \chi)$, and $\mathcal{C}$ is a coherent category, then $[\![\vec{x}\,.\,\phi]\!]$ is the union of the subobjects $[\![\vec{x}\,.\,\psi]\!]$ and $[\![\vec{x}\,.\,\chi]\!]$.

(vii) If ϕ is $(\psi \Rightarrow \chi)$, and $\mathcal{C}$ is a Heyting category, then $[\![\vec{x}\,.\,\phi]\!]$ is the implication $[\![\vec{x}\,.\,\psi]\!] \Rightarrow [\![\vec{x}\,.\,\chi]\!]$ in the Heyting algebra $\mathrm{Sub}(M(A_1,\ldots,A_n))$ (cf. A1.4.13).

(viii) Similarly, if ϕ is $\neg\psi$, and $\mathcal{C}$ is a Heyting category, then $[\![\vec{x}\,.\,\phi]\!]$ is the Heyting negation $\neg[\![\vec{x}\,.\,\psi]\!]$.

(ix) If ϕ is $(\exists y)\psi$ where y is of sort B, and $\mathcal{C}$ is a regular category, then $[\![\vec{x}\,.\,\phi]\!]$ is the image of the composite

$$[\![\vec{x}, y\,.\,\psi]\!] \rightarrowtail M(A_1,\ldots,A_n,B) \xrightarrow{\pi} M(A_1,\ldots,A_n),$$

where π is the projection on the first n factors.

(x) If ϕ is $(\forall y)\psi$ where y is of sort B, and $\mathcal{C}$ is a Heyting category, then $[\![\vec{x}\,.\,\phi]\!]$ is $\forall_\pi([\![\vec{x}, y\,.\,\psi]\!])$, where π is the same projection as in (ix).

(xi) If ϕ is $\bigvee_{i\in I}\psi_i$, and $\mathcal{C}$ is a geometric category, then $[\![\vec{x}\,.\,\phi]\!]$ is the union of the $[\![\vec{x}\,.\,\psi_i]\!]$ in $\mathrm{Sub}(M(A_1,\ldots,A_n))$.

(xii) If ϕ is $\bigwedge_{i\in I}\psi_i$, and $\mathcal{C}$ has arbitrary intersections of subobjects, then $[\![\vec{x}\,.\,\phi]\!]$ is the intersection of the $[\![\vec{x}\,.\,\psi_i]\!]$.

Examining the clauses of the above definition, we see that if M is a Σ-structure in a cartesian category $\mathcal{C}$, we can assign interpretations in M to all Horn formulae-in-context over Σ (though, as we shall see in the next section, this is not best possible); and if $\mathcal{C}$ is a regular (resp. coherent, Heyting, geometric) category then we can assign interpretations to all regular (resp. coherent, first-order, geometric) formulae-in-context over Σ. (In fact, in a geometric category, we can assign interpretations to all infinitary first-order formulae, since we know from A1.4.18 that a geometric category is a Heyting category and that its subobject lattices have arbitrary intersections. However, geometric functors (that is, regular functors preserving arbitrary unions of subobjects – note that the

inverse image functors of geometric morphisms are examples of such functors) do not in general preserve this extra structure, and so they will not preserve the interpretations of non-geometric formulae.) To avoid tedious repetition in what follows, we shall use the phrase 'ϕ is a formula interpretable in $\mathcal{C}$' to mean that $\mathcal{C}$ belongs to one of the classes of categories mentioned above, and ϕ is in the corresponding class of formulae. (We shall similarly speak of sequents and theories being interpretable in $\mathcal{C}$.)

When we introduced the notions of regular, coherent and Heyting categories in Chapter A1, we were at pains to ensure that the appropriate structure (images, unions, universal quantification) should not only exist in the categories in question, but should be stable under pullback. (In the first two cases this was specifically included as part of the definition, but for universal quantification it was a consequence of the other assumptions made by that stage – see A1.4.11.) Thus the Substitution Property has an immediate extension from terms to formulae, which is proved by a similarly straightforward induction:

Lemma 1.2.7 *Let $\vec{y}.\phi$ be a formula-in-context over Σ interpretable in $\mathcal{C}$; let $\vec{s}$ be a string of terms of the same length and type as $\vec{y}$, and let $\vec{x}$ be a context suitable for all the terms in $\vec{s}$. Then, for any Σ-structure M in $\mathcal{C}$, there is a pullback square*

$$\begin{tikzcd}
[\![\vec{x}.\phi[\vec{s}/\vec{y}]]\!]_M \arrow[rr] \arrow[d, tail] & & [\![\vec{y}.\phi]\!]_M \arrow[d, tail] \\
M(A_1,\ldots,A_n) \arrow[rr, "([\![\vec{x}.s_1]\!]_M,\ldots,[\![\vec{x}.s_m]\!]_M)"] & & M(B_1,\ldots,B_m)
\end{tikzcd}$$

where the A_i and B_j are the sorts of the variables x_i and y_j. □

In particular, it follows that knowledge of the interpretation of ϕ in its canonical context determines its interpretation in any other suitable context, by pullback along an appropriate product projection. As with terms, we sometimes abbreviate $[\![\vec{x}.\phi]\!]$ to $[\![\phi]\!]$ if $\vec{x}$ is the canonical context for ϕ. Another easy consequence of 1.2.7 gives a sufficient condition for discarding quantifiers over 'non-existent' variables:

Corollary 1.2.8

(i) *Let $\mathcal{C}$ be (at least) a regular category; let $\vec{x}.\phi$ be a formula-in-context over Σ interpretable in $\mathcal{C}$, and let $y\colon B$ be a variable not appearing in $\vec{x}$ (and hence not free in ϕ). Then, for any Σ-structure M in $\mathcal{C}$ such that the interpretation MB of B is well-supported (i.e. such that $MB \to 1$ is a cover), the interpretations $[\![\vec{x}.\phi]\!]_M$ and $[\![\vec{x}.(\exists y)\phi]\!]_M$ are equal.*

(ii) *If $\mathcal{C}$ is a Heyting category and the hypotheses of* (i) *are otherwise unaltered, then the same conclusion holds for $[\![\vec{x}.\phi]\!]_M$ and $[\![\vec{x}.(\forall y)\phi]\!]_M$.*

Proof (i) If $MB \to 1$ is a cover, then so are the bottom and top edges of the pullback square

$$\begin{tikzcd} \llbracket \vec{x}, y \,.\, \phi \rrbracket_M \arrow[r, two heads] \arrow[d, tail] & \llbracket \vec{x} \,.\, \phi \rrbracket_M \arrow[d, tail] \\ M(A_1, \ldots, A_n, B) \arrow[r, two heads, "\pi"] & M(A_1, \ldots, A_n) \end{tikzcd}$$

(where π is projection on the first n factors), and so the top and right edges form the image factorization of the diagonal.

(ii) In (i), we have shown that the composite $\exists_\pi \circ \pi^*$ is the identity on $\mathrm{Sub}(M(A_1, \ldots, A_n))$. Hence so is its right adjoint $\forall_\pi \circ \pi^*$. □

We note that if there exists a closed term of sort B, then the interpretation of B in any Σ-structure will be well-supported, since the interpretation (in the empty context) of any such term will be a splitting for $MB \to 1$.

When we turn to the analogue for formulae of 1.2.5, life becomes more complicated: not all formulae have interpretations which are 'natural' with respect to arbitrary homomorphisms of Σ-structures. The best we can do is

Lemma 1.2.9 *Let $\mathcal{C}$ be (at least) a cartesian category, let $h\colon M \to N$ be a homomorphism of Σ-structures in $\mathcal{C}$, and let $\vec{x} \,.\, \phi$ be a formula-in-context which is both geometric and interpretable in $\mathcal{C}$. Then there is a commutative square*

$$\begin{tikzcd} \llbracket \vec{x} \,.\, \phi \rrbracket_M \arrow[r, tail] \arrow[d] & M(A_1, \ldots, A_n) \arrow[d, "h_{A_1} \times \cdots \times h_{A_n}"] \\ \llbracket \vec{x} \,.\, \phi \rrbracket_N \arrow[r, tail] & N(A_1, \ldots, A_n) \end{tikzcd} .$$

Proof This is once again an induction over the structure of ϕ, using the fact that the diagram commutes iff $\llbracket \vec{x} \,.\, \phi \rrbracket_M$ is contained in the pullback of $\llbracket \vec{x} \,.\, \phi \rrbracket_N$ along the right vertical map. □

One reason why 1.2.9 does not extend to more general first-order formulae is that implication is not order-preserving in its first variable: if $X_1 \leq Y_1$ and $X_2 \leq Y_2$ in some subobject lattice in a Heyting category, we do not necessarily have $(X_1 \Rightarrow X_2) \leq (Y_1 \Rightarrow Y_2)$. (Similar problems arise with negation.) These problems can be avoided by requiring that the 'naturality square' of 1.2.9 should not only commute but also be a pullback. But we also have problems with universal quantification: for example, the centre of a group (the interpretation of the formula-in-context $x \,.\, (\forall y)(xy = yx)$) is not 'natural' with respect to group homomorphisms. When dealing with first-order theories, we therefore have to

build the naturality result for arbitrary formulae into our notion of morphism of structures:

Definition 1.2.10 (*a*) We call a homomorphism $h\colon M \to N$ of Σ-structures in a Heyting category an *elementary morphism* if, for each first-order formula-in-context $\vec{x}.\phi$ over Σ, there is a commutative square

$$\begin{tikzcd} [\![\vec{x}.\phi]\!]_M \arrow[r, tail] \arrow[d] & M(A_1,\dots,A_n) \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] \\ [\![\vec{x}.\phi]\!]_N \arrow[r, tail] & N(A_1,\dots,A_n) \end{tikzcd} .$$

(Note that, thanks to 1.2.7, it is sufficient to demand this condition only for formulae in their canonical contexts.)

(*b*) We call h an *elementary embedding* if the square above not only commutes but is a pullback for all first-order formulae-in-context.

(*c*) We call h an *embedding* (or we say that M is a *substructure* of N) if the square above is a pullback for all atomic formulae (including equality). (Note that this requirement forces each h_A to be monic, since the naturality square

$$\begin{tikzcd} MA \arrow[r, tail, "\Delta"] \arrow[d, "h_A"] & MA \times MA \arrow[d, "h_A \times h_A"] \\ NA \arrow[r, tail, "\Delta"] & NA \times NA \end{tikzcd}$$

for the equality relation of sort A is a pullback. Also, a straightforward inductive argument ensures that the naturality square is then a pullback for all quantifier-free formulae-in-context.)

For structures in a Boolean coherent category, the notions of 'elementary morphism' and 'elementary embedding' coincide, since the commutativity of the squares of 1.2.10(*a*) for two complementary formulae ϕ and $\neg\phi$ forces both of them to be pullbacks (cf. the proof of A1.6.5). However, in the non-Boolean case an elementary morphism does not even have to be monic, as the following example shows:

Example 1.2.11 Let $\mathcal{C}$ be the functor category $[\mathbf{2}, \mathbf{Set}]$ (where $\mathbf{2}$ is the two-element ordered set $\{0, 1\}$), and let Σ be the signature with one sort and no primitive symbols except equality (so that Σ-structures are simply objects). Let M be any object of $\mathcal{C}$ such that $M(1)$ is a singleton, and let N be the terminal object of $\mathcal{C}$. We claim that the unique morphism $h\colon M \to N$ is elementary. To see this, note first that, by an easy inductive argument, the interpretation in N

of any first-order formula-in-context $x_1, \ldots, x_n . \phi$ is either the whole of $N^n \cong 1$ or its zero subobject. Thus, to verify that we have a commutative square as in 1.2.10(a) for all $\vec{x} . \phi$, it suffices to verify that $[\![\vec{x} . \phi]\!]_M = 0$ whenever $[\![\vec{x} . \phi]\!]_N = 0$. But the functor 'evaluate at 1' from $\mathcal{C}$ to **Set** is a Heyting functor (because it is the inverse image of an open inclusion, cf. A4.5.1), and it maps both M and N to singleton sets, so $[\![\vec{x} . \phi]\!]_N = 0$ forces $[\![\vec{x} . \phi]\!]_M(1) = [\![\vec{x} . \phi]\!]_N(1) = \emptyset$. And this in turn forces $[\![\vec{x} . \phi]\!]_M(0)$ to be empty as well.

We now come at last to the formal definition of the validity of an axiom in a structure.

Definition 1.2.12 Let M be a Σ-structure in a category $\mathcal{C}$.

(a) If $\sigma = (\phi \vdash_{\vec{x}} \psi)$ is a sequent over Σ interpretable in $\mathcal{C}$, we say σ is *satisfied in M* (and write $M \models \sigma$) if $[\![\vec{x} . \phi]\!]_M \leq [\![\vec{x} . \psi]\!]_M$ in $\mathrm{Sub}(M(A_1, \ldots, A_n))$.

(b) If $\mathbb{T}$ is a theory over Σ interpretable in $\mathcal{C}$, we say M is a *model* of $\mathbb{T}$ (and write $M \models \mathbb{T}$) if all the axioms of $\mathbb{T}$ are satisfied in M.

(c) We write $\mathbb{T}$-$\mathbf{Mod}(\mathcal{C})$ for the full subcategory of Σ-$\mathbf{Str}(\mathcal{C})$ whose objects are models of $\mathbb{T}$, and $\mathbb{T}$-$\mathbf{Mod}(\mathcal{C})_e$ for the non-full subcategory with the same objects but only elementary morphisms.

Lemma 1.2.13 *Let $T: \mathcal{C} \to \mathcal{D}$ be a cartesian (resp. regular, coherent, Heyting, geometric) functor between categories of the appropriate kind; let M be a Σ-structure in $\mathcal{C}$, and let σ be a sequent over Σ interpretable in $\mathcal{C}$. If $M \models \sigma$ in $\mathcal{C}$, then Σ-$\mathbf{Str}(T)(M) \models \sigma$ in $\mathcal{D}$. The converse implication holds if T is conservative.*

Proof An easy induction shows that T preserves the interpretations of all formulae-in-context interpretable in the appropriate class of categories; so the first assertion is immediate. For the second, note that a sequent $(\phi \vdash_{\vec{x}} \psi)$ is satisfied in M iff the inclusion $[\![\vec{x} . \phi \wedge \psi]\!]_M \rightarrowtail [\![\vec{x} . \phi]\!]_M$ is an isomorphism. □

In particular, if $\mathbb{T}$ is a regular (coherent, ...) theory over Σ, then for any regular (coherent, ...) functor $T: \mathcal{C} \to \mathcal{D}$ the functor Σ-$\mathbf{Str}(T)$ which we defined after 1.2.1 restricts to a functor $\mathbb{T}$-$\mathbf{Mod}(T): \mathbb{T}$-$\mathbf{Mod}(\mathcal{C}) \to \mathbb{T}$-$\mathbf{Mod}(\mathcal{D})$. (In practice, we usually denote this functor simply by T.) Note also that any such T preserves pullbacks, and so further restricts to a functor between the non-full subcategories of $\mathbb{T}$-models and (elementary) embeddings.

The importance of the second assertion of 1.2.13 is that, in suitable cases, it allows us to reduce the problem of determining which Σ-structures are $\mathbb{T}$-models in an unfamiliar category $\mathcal{C}$ to the corresponding problem in a more familiar category $\mathcal{D}$, by constructing a suitable structure-preserving embedding of $\mathcal{C}$ in $\mathcal{D}$. We give two examples of this process:

Corollary 1.2.14 *Let $\mathbb{T}$ be a geometric theory over a signature Σ. Then*

(i) *For any small category $\mathcal{C}$, a Σ-structure M in $[\mathcal{C}, \mathbf{Set}]$ is a $\mathbb{T}$-model iff each $\mathrm{ev}_c(M)$, $c \in \mathrm{ob}\, \mathcal{C}$, is a $\mathbb{T}$-model in $\mathbf{Set}$, where ev_c denotes the functor*

'evaluate at c'. In fact we have an isomorphism $\mathbb{T}$-**Mod**$([\mathcal{C}, \mathbf{Set}]) \cong [\mathcal{C}, \mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})]$.

(ii) *For any topological space* X, *a* Σ-*structure* M *in* $\mathbf{Sh}(X)$ *is a* $\mathbb{T}$-*model iff each* $x^*(M)$, $x \in X$, *is a* $\mathbb{T}$-*model in* $\mathbf{Set}$, *where* $x^* \colon \mathbf{Sh}(X) \to \mathbf{Set}$ *is the stalk functor associated with* x *(i.e. the inverse image functor along* x *regarded as a continuous map* $1 \to X$, *cf.* C1.3.2*).*

Proof In each case, we have merely to observe that the functors ev_c (resp. x^*) are geometric (because they are inverse image functors), and that they are jointly conservative. For the second assertion of (i), we also need to observe the (obvious) fact that a homomorphism of Σ-structures in $[\mathcal{C}, \mathbf{Set}]$ is essentially the same thing as a natural transformation between functors $\mathcal{C} \to \Sigma\text{-}\mathbf{Str}(\mathbf{Set})$. □

Note that 1.2.14(ii) is valid only for spaces X, and not for more general locales (cf. C1.2.3): indeed the assertion, for a locale X, that the functors x^* are jointly conservative (as x ranges over all points of X) is equivalent to the assertion that X is spatial, as we saw in C1.4.9.

Examples 1.2.15 We conclude this section by revisiting some of the examples of theories considered in 1.1.7, and describing their models in appropriate categories.

(a) If $\mathbb{T}$ is an algebraic theory, then $\mathbb{T}$-models in a cartesian category $\mathcal{C}$ are just what we usually mean by objects equipped with algebraic structure (group objects, R-module objects, ...) in $\mathcal{C}$. The key point in proving this is that a sequent $(\top \vdash_{\vec{x}} (s = t))$ is satisfied in a structure M iff the equalizer of $[\![\vec{x}\,.\,s]\!]_M$ and $[\![\vec{x}\,.\,t]\!]_M$ is the whole of their domain, iff the interpretations of $\vec{x}\,.\,s$ and $\vec{x}\,.\,t$ in M are equal. Similar remarks apply to the many-sorted algebraic case. (When dealing with algebraic theories, we tend to speak of *algebras* rather than models of a theory.)

(c) Again, a model of the theory of posets, as described in 1.1.7(c), in a cartesian category $\mathcal{C}$ is just an internal poset in $\mathcal{C}$, as usually defined. For example, if we write $(a, b) \colon M_1 \rightarrowtail M$ for the interpretation in a Σ-structure M of the relation symbol $\leq$ (and we have suppressed the name of the unique sort of Σ, to simplify the notation), then the third (transitivity) axiom of 1.1.7(c) becomes the assertion that

$$\pi_{12}^*(M_1) \cap \pi_{23}^*(M_1) \leq \pi_{13}^*(M_1)$$

in $\mathrm{Sub}(M \times M \times M)$, where $\pi_{ij} \colon M \times M \times M \to M \times M$ is projection on the ith and jth factors. But this is equivalent to saying that, if we form the pullback

$$\begin{tikzcd} M_2 \arrow[r, "q"] \arrow[d, "p"'] & M_1 \arrow[d, "a"] \\ M_1 \arrow[r, "b"] & M \end{tikzcd} ,$$

the morphism $(ap, bq)\colon M_2 \to M \times M$ factors through $(a, b)\colon M_1 \rightarrowtail M \times M$ – which is how we defined transitivity for relations in a cartesian category in A1.3.6(d).

(e) If $\mathcal{C}$ is a regular category, then a model of the theory of categories in $\mathcal{C}$ is, yet again, equivalent to an internal category in $\mathcal{C}$ as defined in B2.3.1. But the latter definition is interpretable in any cartesian category, whether or not it is regular. The reason is as follows: if (adopting the notation of B2.3.1) we write C_0 and C_1 for the objects interpreting the sorts O and M respectively, and $C_2 \rightarrowtail C_1 \times C_1 \times C_1$ for the interpretation of the ternary relation symbol T, then the existential quantifier in the third displayed sequent of 1.1.7(e) tells us to form the image of the composite

$$C_2 \rightarrowtail C_1 \times C_1 \times C_1 \xrightarrow{\pi_{12}} C_1 \times C_1;$$

but the validity of the second displayed sequent tells us precisely that this composite is monic, and so it has an image factorization (which is, moreover, stable under pullback) even if arbitrary morphisms of $\mathcal{C}$ do not. The idea underlying this, that we are allowed to form images of morphisms which are already known to be monic, will be the key ingredient in the definition of a cartesian theory in the next section.

(f) If A is an abelian group object in a regular category $\mathcal{C}$, the divisibility axioms of 1.1.7(f) are satisfied in A precisely if the nth power map $[\![x \,.\, nx]\!]\colon A \to A$ is a cover, for all $n > 1$. Once again, we note that the torsion-freeness axioms of 1.1.7(d) (plus a little abelian group theory) assert that these maps are monic; so the conjunction of the two says that they are isomorphisms, which makes sense in any cartesian category.

(g) We shall not attempt to describe what it means to be a model of the theory of local rings in an arbitrary coherent category; but we note that for the category $\mathbf{Sh}(X)$ of sheaves on a topological space X, it reduces by 1.2.14(ii) to the notion of a sheaf of rings A on X (that is, a ring object in $\mathbf{Sh}(X)$) whose stalks are all local rings – and this is what algebraic geometers usually mean when they talk about a 'sheaf of local rings'. Note that this condition does not imply that the ring $A(U)$ of sections of such a ring object over an open set U is local: the functor 'evaluate at U' is (cartesian but) not coherent. Even in the simple case when X is a discrete two-point space $\{x, y\}$, and U is the whole of X, we can observe this failure: $A(X)$ is then the cartesian product of the stalks $x^*(A)$ and $y^*(A)$, and a cartesian product of two local rings is not local. Similar remarks apply to the (coherent) theory of fields, described in 1.1.7(h), and to the geometric theory of torsion groups described in 1.1.7(j).

(k) Again, we shall not describe the models of the theory ($\mathbb{K}$, say) of 1.1.7(k) in a general geometric category, but it is instructive to note that its models in a Grothendieck topos $\mathcal{E}$ are exactly the Kuratowski-finite objects of $\mathcal{E}$, as defined in Section D5.4 below. (The proof of this fact is essentially contained in 5.4.13.) Note, however, that morphisms of $\mathbb{K}$-models are epimorphisms rather

than arbitrary morphisms between K-finite objects, since they are required to preserve the validity of the relations R_n.

(m) Finally, we discuss the propositional theory $\mathbb{P}$ of completely prime filters in a frame L, introduced in 1.1.7(m). Clearly, a structure for the appropriate signature in a geometric category $\mathcal{C}$ is simply a function assigning to each atomic proposition F_a a subterminal object $U_a \rightarrowtail 1$ in $\mathcal{C}$; and the assertion that such a structure is a model of $\mathbb{P}$ is equivalent to saying that $a \mapsto U_a$ preserves finite meets and arbitrary joins, i.e. that it is a frame homomorphism. In particular, if $\mathcal{C} = \mathbf{Sh}(X)$ is the topos of sheaves on a locale X, then the lattice $\mathrm{Sub}_{\mathcal{C}}(1)$ is isomorphic to $\mathcal{O}(X)$, by C1.3.15; so $\mathbb{P}$-models in $\mathbf{Sh}(X)$ correspond to locale maps $X \to Y$, where Y is the locale defined by $\mathcal{O}(Y) = L$. And by C1.4.5 these in turn correspond to (isomorphism classes of) geometric morphisms $\mathbf{Sh}(X) \to \mathbf{Sh}(Y)$. Moreover, a homomorphism of models in this case is just the same thing as an inequality between morphisms in $\mathbf{Loc}$, as defined in C1.1.6; so we actually have an equivalence of categories

$$\mathbb{P}\text{-}\mathbf{Mod}(\mathbf{Sh}(X)) \simeq \mathfrak{Top}\,(\mathbf{Sh}(X), \mathbf{Sh}(Y)) .$$

This is a part of the assertion that $\mathbf{Sh}(Y)$ is a *classifying topos* for the theory $\mathbb{P}$; see 3.1.14 below.

Suggestions for further reading: Kock & Reyes [640], Makkai & Reyes [790], Reyes [1004].

D1.3 First-order logic

The definitions of the last two sections provide an extremely useful tool for constructing objects and morphisms with prescribed properties in a given category $\mathcal{C}$. Starting with (a structure for) a signature which names some 'basic' objects and morphisms of $\mathcal{C}$, we can use the terms and formulae over this signature to name more complicated objects and morphisms via the categorical semantics developed in the last section.

But first-order logic is more than a convenient shorthand for *describing* particular objects and morphisms of a category; it is also a tool for *proving* things about them via a suitable deduction-system. Our first concern in this section is to develop such deduction-systems for the various fragments of logic which we have considered, and to prove that they are *sound* for the categorical semantics; i.e. that anything which is formally derivable in the deduction-system will be valid in any structure for the given signature in any category of the appropriate type. Of course, we should also like the deduction-systems to be *complete* in the sense that the converse of the last assertion holds: anything which is valid in all structures in all suitable categories should be formally derivable in the logic. The completeness theorem is indeed true for the deduction-systems we shall construct, but the proof of this result will be deferred to Section D1.4.

Our deduction-systems will be formulated as *sequent calculi*; that is, they will provide rules for inferring the validity of certain sequents (as defined in 1.1.5) from certain other sequents. (There are other possible ways of presenting them, but experience suggests that a sequent calculus is the most convenient for our purposes.) We shall write the rules in the form

$$\frac{\Gamma}{\sigma}$$

where Γ is a (possibly empty) list of sequents and σ is a sequent; the interpretation of the rule being that if we have established the validity of all the sequents in Γ we may infer the validity of σ. In the particular case when Γ is empty, we shall say that σ is a (logical) *axiom*, and omit the line above it. We shall also use the shorthand

$$\frac{\sigma}{\overline{\tau}}$$

to indicate that both $\frac{\sigma}{\tau}$ and $\frac{\tau}{\sigma}$ are rules of inference, i.e. that either of σ and τ may be inferred from the other. A *derivation* in the deduction-system will then have the form of a tree such as

$$\dfrac{\dfrac{\sigma_1 \qquad \sigma_2}{\sigma_3} \qquad \sigma_4}{\sigma_5}$$

where each sequent without a line above it is an axiom, and each sequent below a line follows by a rule of inference from those above the line. (If our language includes either of the infinitary connectives $\bigvee$ and $\bigwedge$, the tree may be infinite; but we still require it to be *well-founded* in the sense that it has no infinite ascending chains.) A *derivation relative to a theory* $\mathbb{T}$ (defined as in 1.1.6) will be a similar tree, except that the sequents without lines above them are allowed to include (non-logical) axioms of $\mathbb{T}$ as well as the logical axioms to be listed below.

Apart from the 'structural' rules for manipulating sequents, and the rules for handling the equality predicate, the rules mostly have the form of 'introduction' or 'elimination' rules for the various logical connectives and quantifiers. Of course, in each fragment of logic which we consider, we adopt the introduction and elimination rules for those connectives and quantifiers which are present in the language (as well as the structural and equality rules, which are present in all fragments); thus, for example, *coherent logic* is specified by the rules in parts (a), (b), (c), (d), (f) and (i) of the following definition, and *full (finitary) first-order logic* is obtained by adding the rules in (e) and (g) (and deleting (i)). Throughout the definition, it is assumed that all the sequents which appear are well-formed, i.e. that the contexts which appear in them are suitable for the formulae on either side; but, apart from this restriction, the formulae ϕ, ψ and χ are allowed to be anything in the appropriate fragment of the first-order language.

Definition 1.3.1 (*a*) The *structural rules* consist of the *identity axiom*

$$(\phi \vdash_{\vec{x}} \phi),$$

the *substitution rule*

$$\frac{(\phi \vdash_{\vec{x}} \psi)}{(\phi[\vec{s}/\vec{x}] \vdash_{\vec{y}} \psi[\vec{s}/\vec{x}])}$$

where $\vec{y}$ is any string of variables including all the variables occurring in the string of terms $\vec{s}$, and the *cut rule*

$$\frac{(\phi \vdash_{\vec{x}} \psi) \ (\psi \vdash_{\vec{x}} \chi)}{(\phi \vdash_{\vec{x}} \chi)} .$$

Note that the substitution rule allows the possibility of making the trivial substitution $[\vec{x}/\vec{x}]$; thus it includes the *weakening rule* which says that we may derive $(\phi \vdash_{\vec{y}} \psi)$ from $(\phi \vdash_{\vec{x}} \psi)$ provided $\vec{y}$ contains all the variables in $\vec{x}$. Note also that the converse of the weakening rule is *not* a permitted rule of inference; this is particularly important when applying the cut rule, as it means that we cannot 'eliminate' any variables which may occur in ψ but not in ϕ or χ from the context of the conclusion of this rule. (However, we *can* eliminate a non-occurring variable from a context if it contains another variable of the same sort, by substituting the latter for the former; and similarly if there is a closed term of the appropriate sort.)

(*b*) The *equality rules* consist of the axioms

$$(\top \vdash_{x} (x = x))$$

and

$$(((\vec{x} = \vec{y}) \wedge \phi) \vdash_{\vec{z}} \phi[\vec{y}/\vec{x}])$$

where $\vec{x}$ and $\vec{y}$ are contexts of the same length and type, $(\vec{x} = \vec{y})$ is a shorthand for $((x_1 = y_1) \wedge \cdots \wedge (x_n = y_n))$, and $\vec{z}$ is any context containing $\vec{x}, \vec{y}$ and the free variables of ϕ. From these (plus the structural and conjunction rules), it is a standard exercise to derive the sequents $((x = y) \vdash_{x,y} (y = x))$ and $(((x = y) \wedge (y = z)) \vdash_{x,y,z} (x = z))$, which say that equality is symmetric and transitive; so we could have assumed these as additional axioms if we wished.

(*c*) The *rules for (finite) conjunction* are the axioms

$$(\phi \vdash_{\vec{x}} \top), \quad ((\phi \wedge \psi) \vdash_{\vec{x}} \phi), \quad ((\phi \wedge \psi) \vdash_{\vec{x}} \psi)$$

and the rule

$$\frac{(\phi \vdash_{\vec{x}} \psi) \ (\phi \vdash_{\vec{x}} \chi)}{(\phi \vdash_{\vec{x}} (\psi \wedge \chi))} .$$

(*d*) The *rules for (finite) disjunction* are the 'duals' of those in (*c*): the axioms $(\perp \vdash_{\vec{x}} \phi)$, $(\phi \vdash_{\vec{x}} (\phi \vee \psi))$ and $(\psi \vdash_{\vec{x}} (\phi \vee \psi))$, and the rule

$$\frac{(\phi \vdash_{\vec{x}} \chi) \ (\psi \vdash_{\vec{x}} \chi)}{((\phi \vee \psi) \vdash_{\vec{x}} \chi)} .$$

(*e*) The *rules for implication* consist of the double rule

$$\frac{((\phi \wedge \psi) \vdash_{\vec{x}} \chi)}{\overline{(\psi \vdash_{\vec{x}} (\phi \Rightarrow \chi))}} \,.$$

The *rules for negation* are the special case of this double rule obtained by setting $\chi = \bot$, and identifying $\neg\phi$ with $(\phi \Rightarrow \bot)$.

(*f*) The *rules for existential quantification* consist of the double rule

$$\frac{(\phi \vdash_{\vec{x},y} \psi)}{\overline{((\exists y)\phi \vdash_{\vec{x}} \psi)}}$$

(here our standing hypothesis that all sequents are well-formed includes the information that y is not free in ψ).

(*g*) The *rules for universal quantification* consist of the double rule

$$\frac{(\phi \vdash_{\vec{x},y} \psi)}{\overline{(\phi \vdash_{\vec{x}} (\forall y)\psi)}} \,.$$

(*h*) The *rules for infinitary conjunction and disjunction* are the obvious infinitary analogues of those in (*c*) and (*d*).

(*i*) Finally, we mention two 'mixed axioms' which are required for coherent logic: the *distributive axiom* which links $\wedge$ and $\vee$:

$$((\phi \wedge (\psi \vee \chi)) \vdash_{\vec{x}} ((\phi \wedge \psi) \vee (\phi \wedge \chi)))$$

and the *Frobenius axiom* which links $\wedge$ and $\exists$:

$$((\phi \wedge (\exists y)\psi) \vdash_{\vec{x}} (\exists y)(\phi \wedge \psi))$$

where y is a variable not in the context $\vec{x}$ (and hence not free in ϕ). Both of these sequents are provable in full first-order logic using the implication rules; for example, here is a derivation of the second one:

$$\begin{array}{c}
((\exists y)(\phi \wedge \psi) \vdash_{\vec{x}} (\exists y)(\phi \wedge \psi)) \\ \hline
((\phi \wedge \psi) \vdash_{\vec{x},y} (\exists y)(\phi \wedge \psi)) \\ \hline
(\psi \vdash_{\vec{x},y} (\phi \Rightarrow (\exists y)(\phi \wedge \psi))) \\ \hline
((\exists y)\psi \vdash_{\vec{x}} (\phi \Rightarrow (\exists y)(\phi \wedge \psi))) \\ \hline
((\phi \wedge (\exists y)\psi) \vdash_{\vec{x}} (\exists y)(\phi \wedge \psi)) \,.
\end{array}$$

However, when we are working in fragments of first-order logic which do not include implication, these sequents are not derivable – for the distributive axiom, this may be seen from the fact that the other axioms involving $\wedge$ and $\vee$ are sound for interpretations in arbitrary (not necessarily distributive) lattices. So we must add them both as axioms in coherent logic, and the Frobenius axiom alone in

regular logic. In geometric logic, we similarly add the appropriate infinitary generalization of the distributive axiom, together with the Frobenius axiom. (Incidentally, the converses of both axioms are derivable in coherent logic: for example,

$$\dfrac{\dfrac{((\phi \wedge \psi) \vdash_{\vec{x},y} \phi)}{((\exists y)(\phi \wedge \psi) \vdash_{\vec{x}} \phi)} \qquad \dfrac{\dfrac{((\phi \wedge \psi) \vdash_{\vec{x},y} \psi) \quad \dfrac{((\exists y)\psi \vdash_{\vec{x}} (\exists y)\psi)}{(\psi \vdash_{\vec{x},y} (\exists y)\psi)}}{((\phi \wedge \psi) \vdash_{\vec{x},y} (\exists y)\psi)}}{((\exists y)(\phi \wedge \psi) \vdash_{\vec{x}} (\exists y)\psi)}}{((\exists y)(\phi \wedge \psi) \vdash_{\vec{x}} (\phi \wedge (\exists y)\psi))}$$

is a derivation of the converse of Frobenius using only rules in (a), (c) and (f).)

We say a sequent σ is *derivable* or *provable* in a (Horn, regular, coherent, ...) theory $\mathbb{T}$ if there exists a derivation relative to $\mathbb{T}$, in the appropriate fragment of first-order logic, with σ as its 'bottom line'. The soundness theorem may now be proved by an easy induction over the 'height' of a derivation:

Proposition 1.3.2 (Soundness Theorem) *Let $\mathbb{T}$ be a Horn (respectively regular, coherent, first-order, geometric) theory over a signature Σ, and let M be a model of $\mathbb{T}$ in a cartesian (resp. regular, coherent, Heyting, geometric) category $\mathcal{C}$. If σ is a sequent (in the appropriate fragment of the first-order language over Σ) which is provable in $\mathbb{T}$, then $M \models \sigma$.*

Proof We have merely to show, for each of the rules in 1.3.1, that if M (lives in a category of the appropriate kind and) satisfies the sequent(s) appearing above the line in the rule, then it also satisfies the sequent below the line. But this is trivial in almost every case, from the way in which the interpretation of formulae was defined in the last section. For example, the soundness of the substitution rule follows from Lemma 1.2.7 and the fact that pullback of subobjects along a fixed morphism is order-preserving; and the second equality axiom is the assertion that, for any subobject $A' \rightarrowtail A$ in a cartesian category, we have $\pi_1^*(A') \cap \Delta = \pi_2^*(A') \cap \Delta \leq \pi_2^*(A')$ in $\mathrm{Sub}(A \times A)$. The soundness of the double rule for existential quantification is simply the assertion that existential quantification in a regular category is left adjoint to pullback (cf. A1.3.1(iii); note that since ψ does not contain y as a free variable, its interpretation $[\![\vec{x}, y \,.\, \psi]\!]$ is the pullback of $[\![\vec{x} \,.\, \psi]\!]$ along a product projection, by 1.2.7). Similarly for universal quantification in a Heyting category. Finally, the validity of the Frobenius (resp. distributive) axiom in arbitrary regular (resp. coherent) categories was verified in A1.3.3 (resp. A1.4.2). $\square$

Remark 1.3.3 The reader will have observed that our deduction-system is essentially constructive (or intuitionistic, if you prefer), in that the *Law of*

Excluded Middle

$$(\top \vdash_{\vec{x}} (\phi \vee \neg\phi))$$

does not appear as an axiom (and indeed it is not derivable, since it is satisfied in a structure M iff the interpretation $[\![\vec{x}\,.\,\phi]\!]_M$ is a complemented subobject of $MA_1 \times \cdots \times MA_n$, and we know that our logic is sound for interpretations in categories where not all subobjects are complemented). Of course, we obtain *classical first-order logic* by adding this axiom to the rules of 1.3.1(a–g); it follows from 1.3.2 that this logic is sound for arbitrary structures in Boolean coherent categories (recall that all such categories are in fact Heyting categories, by A1.4.10).

We next fulfil a promise made earlier in this chapter, by introducing the notions of cartesian formula and cartesian theory.

Definition 1.3.4 (a) Let $\mathbb{T}$ be a (regular) theory. The class of *cartesian formulae-in-context* relative to $\mathbb{T}$ is defined as follows: atomic formulae are cartesian (in any context), finite conjunctions of cartesian formulae are cartesian, and $\vec{x}\,.\,(\exists y)\phi$ is cartesian provided $\vec{x}, y\,.\,\phi$ is cartesian and the sequent

$$((\phi \wedge \phi[z/y]) \vdash_{\vec{x},y,z} (y = z))$$

is provable in $\mathbb{T}$. (Here z is a variable of the same sort as y, not occurring in $\vec{x}$.) A sequent $(\phi \vdash_{\vec{x}} \psi)$ is said to be cartesian relative to $\mathbb{T}$ if both ϕ and ψ are cartesian relative to $\mathbb{T}$ in the context $\vec{x}$.

(b) A regular theory $\mathbb{T}$ is said to be *cartesian* if there is a well-founded partial ordering of its axioms, such that each axiom is cartesian relative to the subtheory formed by the axioms which precede it in the ordering.

For example, the theory of categories is cartesian, because its only axiom involving existential quantification is the third displayed sequent of 1.1.7(e), and this is cartesian relative to the theory whose only axiom is the second displayed sequent.

If M is a $\mathbb{T}$-model (where $\mathbb{T}$ is, for the moment, a Horn theory) in a cartesian category $\mathcal{C}$, then any cartesian formula-in-context relative to $\mathbb{T}$ has an interpretation in M: we simply redefine the interpretation $[\![\vec{x}\,.\,(\exists y\colon B)\psi]\!]_M$, as compared with the definition in 1.2.6(ix), to be the composite

$$[\![\vec{x}, y\,.\,\psi]\!] \rightarrowtail M(A_1, \ldots, A_n, B) \xrightarrow{\pi} M(A_1, \ldots, A_n),$$

which we know to be monic, by the soundness theorem for Horn logic in cartesian categories. It is clear that the Substitution Lemma 1.2.7 remains valid for this interpretation, and that the rules of 1.3.1(f) and the Frobenius axiom are sound for it; hence we can recursively define what it means for a Σ-structure in $\mathcal{C}$ to be a model for a cartesian theory $\mathbb{T}$, and we obtain

Proposition 1.3.5 *Let $\mathbb{T}$ be a cartesian theory, and σ a cartesian sequent relative to $\mathbb{T}$. If σ is provable in $\mathbb{T}$ (using the rules of regular logic), then it is satisfied in every $\mathbb{T}$-model in a cartesian category.* $\square$

Remark 1.3.6 Another fragment of first-order logic which is of some interest, although we shall not discuss it in detail here, is *disjunctive logic*: this is the sub-fragment of geometric logic (or, in the finitary case, coherent logic) in which formulae are subject to the same restriction as in 1.3.4 whenever an existential quantifier occurs, and additionally to the restriction that disjunctions must be 'provably disjoint', i.e., before we can write down a disjunction $\bigvee_{i \in I} \phi_i$, we need to know that the sequents $((\phi_i \wedge \phi_j) \vdash \bot)$ are provable for all $i \neq j$. An example of a disjunctive theory which is not cartesian is the theory of fields, as described in 1.1.7(h): in the formula

$$((x = 0) \vee (\exists y)(xy = 1))$$

which occurs in the sequent displayed there, the existential quantifier is justified by the fact that inverses in a (commutative) ring are unique when they exist, and the disjunction is justified by the nontriviality axiom $((0 = 1) \vdash \bot)$. On the other hand, the theory of local rings is not disjunctive, since the disjunction occurring in the displayed axiom of 1.1.7(g) is not provably disjoint. (Finitary) disjunctive logic may be interpreted in any category with finite limits, a strict initial object and disjoint (finite) coproducts which are stable under pullback; such categories are naturally called *(finitary) disjunctive categories* (in the finitary case, they are also sometimes known as *lextensive categories*). For more details about this fragment of logic, see [508].

We may also now introduce the notions of provable equivalence of formulae and of of equivalence of theories; the latter was promised after Definition 1.1.6.

Definition 1.3.7 (a) Let ϕ and ϕ' be two formulae, and $\vec{x}$ a context suitable for both of them. We say that ϕ and ϕ' are *provably equivalent* (relative to a theory $\mathbb{T}$) in the context $\vec{x}$ if the sequents $(\phi \vdash_{\vec{x}} \phi')$ and $(\phi' \vdash_{\vec{x}} \phi)$ are both derivable (relative to $\mathbb{T}$). (We sometimes write $(\phi \dashv\vdash_{\vec{x}} \phi')$ to indicate that both these sequents hold.)

(b) Let $\mathbb{T}$ and $\mathbb{T}'$ be two theories over the same signature Σ. We say $\mathbb{T}$ and $\mathbb{T}'$ are *equivalent* if every axiom of $\mathbb{T}'$ is provable (in the appropriate fragment of logic) in $\mathbb{T}$, and conversely.

It is clear that if $\mathbb{T}$ and $\mathbb{T}'$ are equivalent, then exactly the same sequents are provable in each of them; for if we have a derivation of σ relative to $\mathbb{T}'$, we may convert it into a derivation relative to $\mathbb{T}$ by inserting, above each axiom of $\mathbb{T}'$ which appears, a derivation of that sequent relative to $\mathbb{T}$. Hence also equivalence of theories is an equivalence relation. From now on, when we talk of a theory, we shall tend to think of it as an equivalence class rather than as a particular set of axioms. Note, however, that our notion of equivalence is still

signature-dependent: we cannot say, for example, that the theory of Boolean algebras (that is, complemented distributive lattices) is equivalent to the theory of Boolean rings (that is, rings in which every element x satisfies $x^2 = x$), because they are written over different signatures. Later on (in 1.4.9), we shall introduce the weaker notion of 'Morita equivalence' which frees our theories from this dependence on a particular signature.

Using the notions introduced in 1.3.7, we may simplify the presentations of theories in a way which is often useful.

Lemma 1.3.8
(i) *Any regular formula-in-context is provably equivalent (in the empty theory) to one of the form $(\exists \vec{x})\phi$ where ϕ is a Horn formula (and $(\exists \vec{x})$ is shorthand for $(\exists x_1) \cdots (\exists x_n)$). Moreover, if the original formula is cartesian relative to some theory $\mathbb{T}$, so is the equivalent formula of the above form.*

(ii) *Any coherent formula-in-context is provably equivalent to one of the form $\bigvee_{i=1}^{n} \phi_i$, where the ϕ_i are regular formulae (in the form described in* (i)*). The same holds for geometric formulae, if the disjunction is allowed to be infinitary.*

Proof (i) We observe first that a formula of the form $(\phi \wedge (\exists x)\psi)$ having an existential quantifier 'inside' a conjunction may be replaced by $(\exists x)(\phi \wedge \psi)$ using the Frobenius axiom and its converse, provided x does not occur in ϕ – and we may always assume this is so, since we are free to rename the bound variables of $(\exists x)\psi$. By repeated use of this technique, we may reduce any regular formula to the required form. It is also clear that, if $(\exists x)\psi$ is cartesian relative to $\mathbb{T}$, then so is $(\exists x)(\phi \wedge \psi)$, so we do not destroy cartesianness in this process.

(ii) This time, we first use the distributive axiom and its converse, plus the (easy) fact that $(\exists x)(\phi \vee \psi)$ is provably equivalent to $((\exists x)\phi \vee (\exists x)\psi)$, to 'pull all disjunctions to the front', i.e. to reduce our formula to a disjunction of regular formulae. Then we apply the technique of part (i) to each of the latter. The geometric case is similar. □

It will be noted that, in the above proof, we have made use of the fact that provable equivalence of formulae is 'stable under compounding', i.e. that if ϕ is provably equivalent to ϕ', then $(\phi \wedge \psi)$ is provably equivalent to $(\phi' \wedge \psi)$, and similarly for the other connectives and quantifiers. All such results are easily derived from the rules of inference in 1.3.1.

Remark 1.3.9 We do *not* have any standard form, similar to those of 1.3.8, for first-order formulae-in-context. Indeed, in the constructive logic in which we are working, we cannot even pull all the quantifiers to the front of a first-order formula (as we are accustomed to do in classical logic): the 'dual' of the Frobenius axiom, i.e. the sequent

$$((\forall y)(\phi \vee \psi) \vdash_{\vec{x}} (\phi \vee (\forall y)\psi))$$

where y is not free in ϕ, is not constructively valid, and so the formula on the right is not in general provably equivalent to one with all its quantifiers at the front. (But cf. 5.4.6 below.)

Proposition 1.3.10

(i) *Any Horn theory* $\mathbb{T}$ *is equivalent to one whose axioms have the form* $(\phi \vdash_{\vec{x}} \psi)$, *where* ϕ *is a finite conjunction of atomic formulae (possibly the empty conjunction* $\top$*) and* ψ *is atomic.*

(ii) *Any regular theory (in particular, any cartesian theory)* $\mathbb{T}$ *is equivalent to one whose axioms have the form*

$$(\phi \vdash_{\vec{x}} (\exists \vec{y})\psi)$$

where ϕ *and* ψ *are Horn formulae, and we may additionally assume that the sequent* $(\psi \vdash_{\vec{x},\vec{y}} \phi)$ *is provable (in the empty theory).*

(iii) *Any coherent theory* $\mathbb{T}$ *is equivalent to one whose axioms have the form*

$$(\phi \vdash_{\vec{x}} \bigvee_{i=1}^{m} (\exists \vec{y}_i)\psi_i)$$

where ϕ *and the* ψ_i *are Horn formulae, and the sequents* $(\psi_i \vdash_{\vec{x},\vec{y}_i} \phi)$ *are provable.*

(iv) *Any geometric theory* $\mathbb{T}$ *is equivalent to one whose axioms have the same form as in* (iii), *except that the disjunction on the right may be infinite.*

Proof (i) If we have a sequent of the form $(\phi \vdash \top)$ in the axioms of $\mathbb{T}$, we can leave it out, since such sequents are logical axioms. And if we have a sequent $(\phi \vdash_{\vec{x}} \bigwedge_{i=1}^{n} \psi_i)$ with a nontrivial conjunction on the right, we may replace it by the family of sequents $(\phi \vdash_{\vec{x}} \psi_i)$, $1 \leq i \leq n$. It is straightforward to verify from the rules for finite conjunction (and, in one direction, the cut rule) that this family of sequents is equivalent to the single sequent with which we started.

(ii) We first use Lemma 1.3.8(i) to replace any sequent $(\phi \vdash \psi)$ with an equivalent one in which the formulae on each side have all their existential quantifiers at the front. Then we observe that the rules for existential quantification allow us to remove such quantifiers from the left-hand side of a sequent, at the cost of adding more variables to the context. Finally, having got our axioms into the form $(\phi \vdash_{\vec{x}} (\exists \vec{y}) \,.\, \psi)$ where ϕ and ψ are finite conjunctions of atomic formulae, we may replace the formula ψ on the right by $(\phi \wedge \psi)$ without changing (the equivalence class of) our theory.

(iii) Similarly, we first reduce our axioms to sequents in which the formulae on either side are in the standard form described in 1.3.8(ii). Then we may eliminate any disjunctions on the left-hand side at the cost of increasing the number of axioms (as we did for conjunctions on the right in (i)), and eliminate existential quantifiers on the left as we did in (ii). The final assertion is established by the same trick as in (ii).

(iv) is similar to (iii). □

Given a small category $\mathcal{C}$ with (at least) finite limits, we may define a *canonical signature* $\Sigma_{\mathcal{C}}$ for reasoning about it, as follows. (We shall also make informal use of this notation even when $\mathcal{C}$ is not small, even though we demanded in Definition 1.1.1 that our languages should be 'set-based'; it clearly doesn't make any essential difference if we allow a proper class of sorts.) The signature $\Sigma_{\mathcal{C}}$ has one sort $\ulcorner A \urcorner$ for each object A of $\mathcal{C}$, one function symbol $\ulcorner f \urcorner : \ulcorner A_1 \urcorner \cdots \ulcorner A_n \urcorner \to \ulcorner B \urcorner$ for each morphism $f : A_1 \times \cdots \times A_n \to B$ in $\mathcal{C}$, and one relation symbol $\ulcorner R \urcorner \rightarrowtail \ulcorner A_1 \urcorner \cdots \ulcorner A_n \urcorner$ for each subobject $R \rightarrowtail A_1 \times \cdots \times A_n$ of a finite product in $\mathcal{C}$. The first-order language over this signature is called the *internal language of* $\mathcal{C}$.

There is an evident structure for $\Sigma_{\mathcal{C}}$ in $\mathcal{C}$, which assigns A to $\ulcorner A \urcorner$, f to $\ulcorner f \urcorner$ and R to $\ulcorner R \urcorner$. If this structure satisfies a particular (cartesian, regular, coherent, ...) sequent σ over $\Sigma_{\mathcal{C}}$, we shall say that $\mathcal{C}$ itself satisfies the sequent, and write $\mathcal{C} \models \sigma$. The advantage of this (somewhat tautological) definition is that it allows us to argue 'in the formal language' to prove things about $\mathcal{C}$: for example, if $\mathcal{C}$ is a regular category, then the set of regular sequents over $\Sigma_{\mathcal{C}}$ which are satisfied in $\mathcal{C}$ is closed under the rules of inference in 1.3.1(a–c) and (f), and so anything which we can deduce by these rules from sequents which we know to be valid will also express a true assertion about $\mathcal{C}$. This allows us to reason about $\mathcal{C}$ 'as if its objects were sets and its morphisms were functions' (at least provided we restrict ourselves to regular logic, and argue constructively). Before giving some examples of this process, we first list some instances of how categorical information about $\mathcal{C}$ may be encoded in the formal language.

Lemma 1.3.11 *Let $\mathcal{C}$ be a regular category.*

(i) *$f : A \to A$ is the identity morphism iff $\mathcal{C} \models (\top \vdash_x (\ulcorner f \urcorner(x) = x))$.*

(ii) *$f : A \to C$ is the composite of $g : A \to B$ and $h : B \to C$ iff $\mathcal{C} \models (\top \vdash_x (\ulcorner f \urcorner(x) = \ulcorner h \urcorner(\ulcorner g \urcorner(x))))$.*

(iii) *$f : A \to B$ is monic iff $\mathcal{C} \models ((\ulcorner f \urcorner(x) = \ulcorner f \urcorner(x')) \vdash_x (x = x'))$.*

(iv) *$f : A \to B$ is a cover iff $\mathcal{C} \models (\top \vdash_y (\exists x)(\ulcorner f \urcorner(x) = y))$.*

(v) *A is a terminal object iff $\mathcal{C} \models (\top \vdash_{[]} (\exists x)\top)$ and $\mathcal{C} \models (\top \vdash_{x,x'} (x = x'))$ (here x and x' are of sort $\ulcorner A \urcorner$).*

(vi) *A diagram*

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "h"] \\ C \arrow[r, "k"] & D \end{tikzcd}$$

is a pullback square iff $\mathcal{C}$ satisfies the three sequents

$$(\top \vdash_x (\ulcorner h\urcorner(\ulcorner f\urcorner(x)) = \ulcorner k\urcorner(\ulcorner g\urcorner(x)))),$$
$$(((\ulcorner f\urcorner(x) = \ulcorner f\urcorner(x')) \wedge (\ulcorner g\urcorner(x) = \ulcorner g\urcorner(x'))) \vdash_{x,x'} (x = x')), \text{ and}$$
$$((\ulcorner h\urcorner(y) = \ulcorner k\urcorner(z)) \vdash_{y,z} (\exists x)((\ulcorner f\urcorner(x) = y) \wedge (\ulcorner g\urcorner(x) = z))) .$$

Proof We note that $\mathcal{C} \models (\top \vdash_x (s = t))$ iff the equalizer of $[\![x.s]\!]$ and $[\![x.t]\!]$ is the whole of A (where $\ulcorner A\urcorner$ is the sort of x), iff $[\![x.s]\!]$ and $[\![x.t]\!]$ are the same morphism. Statements (i) and (ii) follow immediately from this. For (iii), we note that by Lemma 1.2.7 the interpretation $[\![x, x'.(\ulcorner f\urcorner(x) = \ulcorner f\urcorner(x'))]\!]$ is the pullback along $f \times f\colon A \times A \to B \times B$ of the diagonal subobject of $B \times B$. This pullback always contains the diagonal subobject of $A \times A$; so the validity of the given sequent is equivalent to saying that

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, tail, "\Delta"] & B \arrow[d, tail, "\Delta"] \\ A \times A \arrow[r, "f \times f"] & B \times B \end{tikzcd}$$

is a pullback, i.e. that f is monic. For (iv), we observe that the interpretation $[\![y.(\exists x)(\ulcorner f\urcorner(x) = y)]\!]$ is the image of the graph of f under the projection $A \times B \to B$; but this is the same as the image of f itself. Now the two sequents in (v) say respectively that the unique morphism $A \to 1$ is a cover and that it is monic; similarly, in (vi), the first sequent says that the given square commutes and the other two say that the canonical morphism from A to the pullback of h and k is an isomorphism. □

From now on, we shall tend to drop the notation $\ulcorner - \urcorner$, and simply identify the sorts, function symbols and relation symbols of the internal language of $\mathcal{C}$ with their interpretations in $\mathcal{C}$. The following 'functional completeness' result is often useful:

Proposition 1.3.12 *Let $\mathcal{C}$ be (at least) a regular category, let A and B be objects of $\mathcal{C}$ and let ϕ be a formula of the internal language of $\mathcal{C}$ having free variables x and y of sorts A and B respectively. Suppose that the sequents $((\phi \wedge \phi[y'/y]) \vdash_{x,y,y'} (y = y'))$ and $(\top \vdash_x (\exists y)\phi)$ are satisfied in $\mathcal{C}$. Then there is a unique morphism $f\colon A \to B$ in $\mathcal{C}$ such that the sequents $(\phi \vdash_{x,y} (f(x) = y))$ and $((f(x) = y) \vdash_{x,y} \phi)$ are valid.*

Proof If we write $(a, b)\colon R \rightarrowtail A \times B$ for the interpretation $[\![x, y.\phi]\!]$, then the two given sequents say respectively that a is monic and that it is a cover. So it

is an isomorphism, and we may take f to be the composite

$$A \xrightarrow{a^{-1}} R \xrightarrow{b} B .$$

It is straightforward to verify that this is the unique morphism with the required property. □

We remark that, although we assumed for simplicity that $\mathcal{C}$ was regular in 1.3.12, we could have obtained the same result for cartesian categories: the second sequent in the statement is cartesian relative to the theory whose only axiom is the first one. The same remark applies to the sequents characterizing finite limits in 1.3.11.

As an application of Proposition 1.3.12, we show how the proofs of (the hard parts of) Propositions A1.3.4 and A1.4.3 can be recast in terms of the internal language. Although these proofs are not different in any essential way from the ones we gave in Chapter A1, they look much more familiar, because they are much more closely modelled on what we should do in order to prove the results in **Set** (that is, we should construct a binary relation and then prove that it is the graph of a function with the required properties).

Lemma 1.3.13 *In a regular category, every cover is a coequalizer.*

Proof Let $f\colon A \twoheadrightarrow B$ be a cover in a regular category $\mathcal{C}$, and let $(a,b)\colon R \rightrightarrows A$ be its kernel-pair. We have to show that f is a coequalizer of a and b. So let $c\colon A \to C$ be any morphism with $ca = cb$, and consider the formula

$$\phi \equiv (\exists x)((f(x) = y) \wedge (c(x) = z))$$

having free variables y, z of sorts B, C respectively. We shall show that ϕ satisfies the sequents of 1.3.12, and so defines a morphism $B \to C$ which is the required factorization of c through f. First we observe that

$$\begin{aligned}(\phi \wedge \phi[z'/z]) &\vdash (\exists x, x')((f(x) = f(x')) \wedge (c(x) = z) \wedge (c(x') = z'))\\ &\vdash (\exists w)((c(a(w)) = z) \wedge (c(b(w)) = z'))\\ &\vdash (z = z'),\end{aligned}$$

where w is of sort R, and the second step uses the characterization of pullbacks in 1.3.11(vi). For the second sequent, we have

$$\begin{aligned}\top &\vdash (\exists x)(f(x) = y)\\ &\vdash (\exists x)((f(x) = y) \wedge (\exists z)(c(x) = z))\\ &\vdash (\exists z)(\exists x)((f(x) = y) \wedge (c(x) = z))\\ &\equiv (\exists z)\phi,\end{aligned}$$

where we have made use of the Frobenius axiom at the third step. □

Lemma 1.3.14 *Let $m_1 : A_1 \rightarrowtail A$ and $m_2 : A_2 \rightarrowtail A$ be subobjects in a coherent category $\mathcal{C}$. Then the square*

$$\begin{tikzcd} A_1 \cap A_2 \arrow[r, tail, "n_1"] \arrow[d, tail, "n_2"'] & A_1 \arrow[d, tail, "p_1"] \\ A_2 \arrow[r, tail, "p_2"] & A_1 \cup A_2 \end{tikzcd}$$

is a pushout in $\mathcal{C}$.

Proof As in the original proof of A1.4.3, we shall simplify notation by writing B, C for $A_1 \cap A_2$ and $A_1 \cup A_2$ respectively. Suppose we are given morphisms $f_i : A_i \to D$ $(i = 1, 2)$ with $f_1 n_1 = f_2 n_2$; so we need to construct a morphism $k : C \to D$ with $kp_i = f_i$ $(i = 1, 2)$. As before, we do so by constructing a suitable coherent formula ϕ in the internal language and applying 1.3.12. Specifically, we take

$$\phi \equiv ((\exists x_1)((p_1(x_1) = z) \wedge (f_1(x_1) = w)) \vee (\exists x_2)((p_2(x_2) = z) \wedge (f_2(x_2) = w))) .$$

Then by the distributive axiom $(\phi \wedge \phi[w'/w])$ entails the disjunction of four formulae, namely

$$(\exists x_i)(\exists x'_j)((p_i(x_i) = p_j(x'_j)) \wedge (f_i(x_i) = w) \wedge (f_j(x'_j) = w'))$$

for $i, j = 1, 2$. For the two cases where $i = j$ we use the fact that p_i is a monomorphism (and 1.3.11(iii)) to conclude that $(x_i = x'_j)$ and hence that $(w = w')$. In the other two cases, we use the fact that the square in the statement is a pullback (and 1.3.11(vi)) to conclude $(\exists y)((n_i(y) = x_i) \wedge (n_j(y) = x'_j))$ and hence $(w = w')$.

To verify that ϕ satisfies the second sequent of 1.3.12 is straightforward, since C is the union of A_1 and A_2 and so we have

$$(\top \vdash_z ((\exists x_1)(p_1(x_1) = z) \vee (\exists x_2)(p_2(x_2) = z)))$$

from which $(\top \vdash_z (\exists w)\phi)$ follows easily.

Hence we have a morphism $k : C \to D$ with $[\![z, w \,.\, \phi]\!] = [\![z, w \,.\, (k(z) = w)]\!]$. Since each p_i is monic, it is easy to deduce that

$$[\![x_i, w \,.\, \phi[p_i(x_i)/z]]\!] = [\![x_i, w \,.\, (f_i(x_i) = w)]\!],$$

and hence that $kp_i = f_i$ $(i = 1, 2)$. The fact that k is the unique morphism with this property follows from the fact that C is the union of its subobjects A_1 and A_2. □

Suggestions for further reading: Bénabou [100], Coste [258], Johnstone [508], Makkai & Reyes [790].

D1.4 Syntactic categories

In 1.3.2 we proved a Soundness Theorem, asserting that 'anything provable is true'; i.e. if a sequent σ is derivable from the axioms of a theory $\mathbb{T}$ (in any of the fragments of logic we have considered), then it is satisfied in all $\mathbb{T}$-models in categories of the appropriate kind. Naturally, we wish to have a converse result asserting that 'anything true is provable'; such a result is known to logicians as a Completeness Theorem. In this section we shall prove such a theorem.

One advantage of considering models in general categories, rather than just in **Set**, is that the theorem has a particularly simple form: given a theory $\mathbb{T}$, we shall construct a category $\mathcal{C}_\mathbb{T}$ of the appropriate kind (the *syntactic category* of $\mathbb{T}$) containing a particular $\mathbb{T}$-model $M_\mathbb{T}$ (the *generic model* of $\mathbb{T}$) with the property that the sequents valid in $M_\mathbb{T}$ are precisely those provable in $\mathbb{T}$. (We shall see in the next section how a more conventional completeness theorem in terms of $\mathbb{T}$-models in **Set** may be extracted from this one.)

The basic idea behind the construction of $\mathcal{C}_\mathbb{T}$ is that its objects are precisely those things which *have* to be present in any category (of the appropriate kind) containing a model of $\mathbb{T}$. More specifically, every object of $\mathcal{C}_\mathbb{T}$ will be the interpretation in $M_\mathbb{T}$ of some formula-in-context of the language of $\mathbb{T}$, and the morphisms of $\mathcal{C}_\mathbb{T}$ will be exactly those morphisms that $\mathbb{T}$ 'tells us' must exist between these objects. For a more formal definition, we shall deal first with the case of a cartesian theory $\mathbb{T}$, and indicate afterwards how things must be modified for more complex theories.

So, for the moment, let $\mathbb{T}$ be a cartesian theory (over a signature Σ, say). We define the objects of $\mathcal{C}_\mathbb{T}$ to be α-equivalence classes of formulae-in-context over Σ which are cartesian relative to $\mathbb{T}$; we shall write $\{\vec{x}\,.\,\phi\}$ for the α-equivalence class of $\vec{x}\,.\,\phi$. Here 'α-equivalence' means something a little broader than it did in Section D1.1: in addition to the bound variables of ϕ, we wish to be free to rename the variables in the context $\vec{x}$, i.e. to regard $\vec{x}\,.\,\phi$ as equivalent to $\vec{y}\,.\,\phi[\vec{y}/\vec{x}]$ if $\vec{y}$ is a context of the same length and type as $\vec{x}$. (In effect, the context itself is regarded as binding the variables which appear in it. Note, incidentally, that if M is any $\mathbb{T}$-model in a cartesian category, then the rules of 1.2.6 ensure that the objects $[\![\vec{x}\,.\,\phi]\!]_M$ and $[\![\vec{y}\,.\,\phi[\vec{y}/\vec{x}]]\!]_M$ are equal (and not merely isomorphic).)

In defining what are the morphisms of $\mathcal{C}_\mathbb{T}$ from $\{\vec{x}\,.\,\phi\}$ to $\{\vec{y}\,.\,\psi\}$, we may thus assume (and it is convenient to do so) that the contexts $\vec{x}$ and $\vec{y}$ are disjoint. Suppose now that we have a cartesian formula θ, whose free variables are all in the context $\vec{x}, \vec{y}$, such that the sequents

$$(\theta \vdash_{\vec{x},\vec{y}} (\phi \wedge \psi)),$$
$$((\theta \wedge \theta[\vec{z}/\vec{y}]) \vdash_{\vec{x},\vec{y},\vec{z}} (\vec{y} = \vec{z})), \text{ and}$$
$$(\phi \vdash_{\vec{x}} (\exists \vec{y})\theta)$$

are provable in $\mathbb{T}$. (Here, as before, $(\vec{y} = \vec{z})$ is shorthand for the conjunction $((y_1 = z_1) \wedge \cdots \wedge (y_m = z_m))$, and $(\exists \vec{y})$ similarly means $(\exists y_1) \cdots (\exists y_m)$. Note, by the way, that the provability of the second sequent ensures that the third is cartesian relative to $\mathbb{T}$.) Then in any $\mathbb{T}$-model M in a cartesian category $\mathcal{C}$, the interpretation $[\![\vec{x}, \vec{y} . \theta]\!]_M$ will (by soundness) be the graph of a morphism from $[\![\vec{x} . \phi]\!]_M$ to $[\![\vec{y} . \psi]\!]_M$. In particular, this must be so in the model $M_{\mathbb{T}}$, so the formula-in-context $\vec{x}, \vec{y} . \theta$ must represent a morphism $\{\vec{x} . \phi\} \to \{\vec{y} . \psi\}$ in $\mathcal{C}_{\mathbb{T}}$. However, it is clear that two formulae θ and θ' must represent the same morphism if they are provably equivalent in $\mathbb{T}$, as defined in 1.3.7(a). (We remark in passing that, if θ and θ' are 'functional' in the sense that the sequents displayed above are provable, then either of the sequents $(\theta \vdash \theta')$ and $(\theta' \vdash \theta)$ implies the other. This is the logical version of an assertion whose categorical formulation may be found at A3.2.3(ii).)

Thus we take the morphisms of $\mathcal{C}_{\mathbb{T}}$ to be $\mathbb{T}$-provable-equivalence classes of formulae-in-context which are $\mathbb{T}$-provably functional, i.e. satisfy the three sequents displayed above. We shall write $[\theta]$ for the provable-equivalence class of θ; we may safely omit its context, since the latter can be inferred from the objects which appear as the domain and codomain of $[\theta]$.

The composite of two morphisms

$$\{\vec{x} . \phi\} \xrightarrow{[\theta]} \{\vec{y} . \psi\} \xrightarrow{[\gamma]} \{\vec{z} . \chi\}$$

is defined to be the $\mathbb{T}$-provable-equivalence class of the formula $(\exists \vec{y})(\gamma \wedge \theta)$; it follows from the remark after 1.3.8 that this does not depend on the choice of the representatives θ and γ, and it is straightforward to verify that it is (cartesian and) $\mathbb{T}$-provably functional from $\{\vec{x} . \phi\}$ to $\{\vec{z} . \chi\}$. It is similarly easy to verify that this composition is associative (up to provable equivalence), and that the provable-equivalence class of

$$\{\vec{x} \,.\, \phi\} \xrightarrow{[(\phi \wedge (\vec{x} = \vec{x'}))]} \{\vec{x'} \,.\, \phi[\vec{x'}/\vec{x}]\}$$

serves as the identity morphism on $\{\vec{x} . \phi\}$ (note the use we have made of α-equivalence; here $\vec{x'}$ is assumed to be a context of the same length and type as $\vec{x}$, but disjoint from it); so we have

Lemma 1.4.1 *$\mathcal{C}_{\mathbb{T}}$ is a category.* □

Next, we prove

Lemma 1.4.2 *$\mathcal{C}_{\mathbb{T}}$ has finite limits.*

Proof The object $\{[\,] . \top\}$ is a terminal object: a formula θ is provably functional from $\{\vec{x} . \phi\}$ to $\{[\,] . \top\}$ iff θ is provably equivalent to ϕ, so there is a unique morphism $\{\vec{x} . \phi\} \to \{[\,] . \top\}$ in $\mathcal{C}_{\mathbb{T}}$. The product of two objects $\{\vec{x} . \phi\}$ and $\{\vec{y} . \psi\}$

(where $\vec{x}$ and $\vec{y}$ are assumed to be disjoint) is $\{\vec{x},\vec{y}.(\phi \wedge \psi)\}$, and the product projections are the provable-equivalence classes

$$\{\vec{x}'.\phi[\vec{x}'/\vec{x}]\} \xleftarrow{[(\phi\wedge\psi\wedge(\vec{x}=\vec{x}'))]} \{\vec{x},\vec{y}.(\phi\wedge\psi)\} \xrightarrow{[(\phi\wedge\psi\wedge(\vec{y}=\vec{y}'))]} \{\vec{y}'.\psi[\vec{y}'/\vec{y}]\}\,.$$

Given morphisms $[\theta]\colon \{\vec{z}.\chi\} \to \{\vec{x}.\phi\}$ and $[\gamma]\colon \{\vec{z}.\chi\} \to \{\vec{y}.\psi\}$, the induced morphism into the product is the provable-equivalence class of $(\theta \wedge \gamma)$; it is again straightforward to verify that this is the unique morphism whose composites with the projections are $[\theta]$ and $[\gamma]$. To form the equalizer of a parallel pair

$$\{\vec{x}.\phi\} \overset{[\theta]}{\underset{[\gamma]}{\rightrightarrows}} \{\vec{y}.\psi\},$$

we take the object $\{\vec{x}'.(\exists\vec{y})(\theta[\vec{x}'/\vec{x}] \wedge \gamma[\vec{x}'/\vec{x}])\}$ and the morphism from this to $\{\vec{x}.\phi\}$ given by the provable-equivalence class of $(\exists\vec{y})(\theta \wedge \gamma \wedge (\vec{x}' = \vec{x}))$. If $[\delta]\colon \{\vec{z}.\chi\} \to \{\vec{x}.\phi\}$ is a morphism having equal composites with $[\theta]$ and $[\gamma]$, then its factorization through the equalizer is simply $[\delta[\vec{x}'/\vec{x}]]$. As before, the verification that all this works is tedious but straightforward. □

Remark 1.4.3 Note that the construction of equalizers in the above proof requires a choice of representatives for the equivalence classes $[\theta]$ and $[\gamma]$; if we make a different choice, we will get a different – though isomorphic – object as the domain of our equalizer. Thus, unless we assume that we have some means of making such a choice – for example, that we have a well-ordering of the formulae in our language, which will be the case if our signature is countable – we cannot assert that $\mathcal{C}_{\mathbb{T}}$ has equalizers in the constructive sense on which we insisted in Section A1.2. The problem could be avoided by factoring the objects of $\mathcal{C}_{\mathbb{T}}$ by provable equivalence as well as α-equivalence; but this tends to become very messy in practice, and it simply means that the arbitrary choices have to be made elsewhere, namely in the proof of 1.4.7 below.

Much as in 1.3.11, properties of objects and morphisms in $\mathcal{C}_{\mathbb{T}}$ are reflected in the $\mathbb{T}$-provability of the appropriate sequents:

Lemma 1.4.4

(i) *A morphism $[\theta]\colon \{\vec{x}.\phi\} \to \{\vec{y}.\psi\}$ is an isomorphism in $\mathcal{C}_{\mathbb{T}}$ iff θ is also $\mathbb{T}$-provably functional from $\{\vec{y}.\psi\}$ to $\{\vec{x}.\phi\}$.*

(ii) *Any object of $\mathcal{C}_{\mathbb{T}}$ is isomorphic to one of the form $\{\vec{x}.\phi\}$ where ϕ is quantifier-free (that is, a finite conjunction of atomic formulae).*

(iii) *$[\theta]$ is a monomorphism in $\mathcal{C}_{\mathbb{T}}$ iff the sequent*

$$((\theta \wedge \theta[\vec{x}'/\vec{x}]) \vdash (\vec{x} = \vec{x}'))$$

is provable in $\mathbb{T}$.

(iv) *Any subobject of* $\{\vec{x}\,.\,\phi\}$ *in* $\mathcal{C}_{\mathbb{T}}$ *is isomorphic to one of the form*

$$\{\vec{x'}\,.\,\psi[\vec{x'}/\vec{x}]\} \xrightarrow{[(\psi \wedge (\vec{x} = \vec{x'}))]} \{\vec{x}\,.\,\phi\}$$

where ψ *is a formula such that the sequent* $(\psi \vdash_{\vec{x}} \phi)$ *is provable in* $\mathbb{T}$. *Moreover, for two such subobjects (corresponding to formulae* ψ *and* χ*), we have* $\{\vec{x}\,.\,\psi\} \leq \{\vec{x}\,.\,\chi\}$ *in* $\mathrm{Sub}(\{\vec{x}\,.\,\phi\})$ *iff the sequent* $(\psi \vdash_{\vec{x}} \chi)$ *is provable in* $\mathbb{T}$.

Proof (i) If θ is provably functional in both directions, then $[\theta]\colon \{\vec{y}\,.\,\psi\} \to \{\vec{x}\,.\,\phi\}$ is easily seen to be inverse to $[\theta]\colon \{\vec{x}\,.\,\phi\} \to \{\vec{y}\,.\,\psi\}$ in $\mathcal{C}_{\mathbb{T}}$. Conversely, if $[\theta]$ has an inverse $[\gamma]$, then one may verify that γ must be $\mathbb{T}$-provably equivalent to θ; so θ is functional in both directions.

(ii) It is clear that if ϕ and ϕ' are provably equivalent formulae in some context $\vec{x}$, then $\{\vec{x}\,.\,\phi\}$ and $\{\vec{x}\,.\,\phi'\}$ are isomorphic objects of $\mathcal{C}_{\mathbb{T}}$; so by 1.3.8(i) we may reduce to the case where ϕ has all its quantifiers at the front. It thus suffices to prove that if $\phi = (\exists y)\psi$ is a cartesian formula (relative to $\mathbb{T}$) in a context $\vec{x}$, then the objects $\{\vec{x}\,.\,\phi\}$ and $\{\vec{x}, y\,.\,\psi\}$ are isomorphic in $\mathcal{C}_{\mathbb{T}}$. But it follows easily from (i) and the fact that ϕ is cartesian that

$$\{\vec{x}, y\,.\,\psi\} \xrightarrow{[(\psi \wedge (\vec{x} = \vec{x'}))]} \{\vec{x'}\,.\,\phi[\vec{x'}/\vec{x}]\}$$

is an isomorphism.

(iii) From the proof of 1.4.2, we may construct the kernel-pair of $[\theta]$ as the object $\{\vec{x}, \vec{x'}\,.\,(\theta \wedge \theta[\vec{x'}/\vec{x}])\}$. Then the provability of the given sequent is equivalent, by (i), to the assertion that the diagonal map from $\{\vec{x}\,.\,\phi\}$ to this object is an isomorphism.

(iv) By (iii), any morphism of the displayed form is clearly monic. But if we are given an arbitrary monomorphism $[\theta]\colon \{\vec{y}\,.\,\psi\} \to \{\vec{x}\,.\,\phi\}$, then (iii) ensures that the formula-in-context $\vec{x}\,.\,(\exists \vec{y})\theta$ is cartesian relative to $\mathbb{T}$, and (i) ensures that $[\theta]$ is an isomorphism (over $\{\vec{x}\,.\,\phi\}$) from $\{\vec{y}\,.\,\psi\}$ to $\{\vec{x}\,.\,(\exists\vec{y})\theta\}$. So every subobject is isomorphic to one of the displayed form. For the final assertion, we have merely to note that $[(\psi \wedge (\vec{x} = \vec{x'}))]\colon \{\vec{x'}\,.\,\psi[\vec{x'}/\vec{x}]\} \to \{\vec{x}\,.\,\chi\}$ is a morphism of $\mathcal{C}_{\mathbb{T}}$ iff $(\psi \vdash \chi)$ is provable in $\mathbb{T}$; and it is the only possible morphism over $\{\vec{x}\,.\,\phi\}$ between these two objects. □

We have a canonical Σ-structure $M_{\mathbb{T}}$ in $\mathcal{C}_{\mathbb{T}}$: it assigns to a sort A the object $\{x\,.\,\top\}$ where x is a variable of sort A, to a function symbol $f\colon A_1 \cdots A_n \to B$ the morphism

$$\{x_1, \ldots, x_n\,.\,\top\} \xrightarrow{[(f(x_1, \ldots, x_n) = y)]} \{y\,.\,\top\}$$

and to a relation symbol $R \rightarrowtail A_1 \cdots A_n$ the subobject of $\{x_1, \ldots, x_n\,.\,\top\}$ whose domain is $\{x_1, \ldots, x_n\,.\,R(x_1, \ldots, x_n)\}$ (as in 1.4.4(iv)). It is now easy to establish

Lemma 1.4.5

(i) *For any term-in-context* $\vec{x}.t$ *over* Σ, *the interpretation* $[\![\vec{x}.t]\!]$ *in* $M_{\mathbb{T}}$ *is the morphism*

$$\{\vec{x}.\top\} \xrightarrow{[(t(\vec{x})=y)]} \{y.\top\}\,.$$

(ii) *For any cartesian formula-in-context* $\vec{x}.\phi$ *over* Σ, *the interpretation* $[\![\vec{x}.\phi]\!]$ *in* $M_{\mathbb{T}}$ *is the subobject* $\{\vec{x}.\phi\} \rightarrowtail \{\vec{x}.\top\}$ *(as in* 1.4.4(iii)*).*

(iii) *A cartesian sequent* $(\phi \vdash_{\vec{x}} \psi)$ *is satisfied in* $M_{\mathbb{T}}$ *iff it is provable in* $\mathbb{T}$.

Proof (i) and (ii) are straightforward inductions over the structure of t and ϕ. (iii) now follows immediately from (ii) and 1.4.4(iv). □

Theorem 1.4.6 (Completeness Theorem for cartesian logic) *Let* $\mathbb{T}$ *be a cartesian theory. If a* $\mathbb{T}$*-cartesian sequent* σ *is satisfied in all models of* $\mathbb{T}$ *in cartesian categories, then it is provable in* $\mathbb{T}$.

Proof From 1.4.5(iii), we deduce in particular that $M_{\mathbb{T}}$ is a model of $\mathbb{T}$. So a sequent satisfied in all $\mathbb{T}$-models must in particular be satisfied in $M_{\mathbb{T}}$, and hence must be provable in $\mathbb{T}$. □

The generic model $M_{\mathbb{T}}$ is more than a convenient tool for proving the completeness theorem, however: it also has an important universal property, namely that all $\mathbb{T}$-models are images of it under cartesian functors. More formally, we have the following result (where we write $\mathfrak{Cart}\,(\mathcal{C},\mathcal{D})$ for the category of cartesian functors from $\mathcal{C}$ to $\mathcal{D}$, and all natural transformations between them):

Theorem 1.4.7 *Let* $\mathbb{T}$ *be a cartesian theory. Then, for any cartesian category* $\mathcal{D}$, *the functor*

$$\mathfrak{Cart}\,(\mathcal{C}_{\mathbb{T}},\mathcal{D}) \longrightarrow \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{D})$$

which sends $F\colon \mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ *to* $F(M_{\mathbb{T}})$ *is (part of) an equivalence of categories.*

Proof We saw in 1.2.13 that the image of a $\mathbb{T}$-model under a cartesian functor is a $\mathbb{T}$-model, so the indicated functor does take values in $\mathbb{T}$-$\mathbf{Mod}(\mathcal{D})$. Conversely, given a $\mathbb{T}$-model M in $\mathcal{D}$, we define $F_M\colon \mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ by setting

$$F_M(\{\vec{x}.\phi\}) = [\![\vec{x}.\phi]\!]_M$$

for each formula-in-context $\vec{x}.\phi$, and taking $F_M([\theta])$ to be the morphism whose graph is the interpretation $[\![\vec{x},\vec{y}.\theta]\!]_M$. (By soundness, this does not depend on the choice of the representative θ.) It is straightforward to verify that F_M is a functor, and (given the description of finite limits in $\mathcal{C}_{\mathbb{T}}$ in the proof of 1.4.2) that it is cartesian. Moreover, if $h\colon M \to N$ is a homomorphism of $\mathbb{T}$-models in $\mathcal{D}$, then 1.2.9 ensures that we have morphisms $[\![\vec{x}.\phi]\!]_M \to [\![\vec{x}.\phi]\!]_N$ for every formula-in-context, and these form a natural transformation $F_M \to F_N$. So $M \mapsto F_M$ becomes a functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{D}) \to \mathfrak{Cart}\,(\mathcal{C}_{\mathbb{T}},\mathcal{D})$.

It is clear that $F_M(M_{\mathbb{T}}) \cong M$, and that this isomorphism is natural in M. Conversely, if we are given any cartesian functor $F: \mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ with $F(M_{\mathbb{T}}) \cong M$, then an easy induction over the structure of ϕ produces isomorphisms $F(\{\vec{x}.\phi\}) \cong [\![\vec{x}.\phi]\!]_M$, which are the components of a natural isomorphism $F \cong F_M$. And this isomorphism is itself natural in F; so we have the required equivalence of categories. □

Thus we may identify models of a cartesian theory $\mathbb{T}$, up to isomorphism, with cartesian functors defined on $\mathcal{C}_{\mathbb{T}}$. In the converse direction, we have

Example 1.4.8 For any small cartesian category $\mathcal{C}$, there is a cartesian theory whose models in an arbitrary cartesian category $\mathcal{D}$ correspond to cartesian functors $\mathcal{C} \to \mathcal{D}$. To construct it, we take the canonical signature $\Sigma_{\mathcal{C}}$, as defined before 1.3.11 (but restricted to unary function symbols, and no relation symbols), and write down all the sequents which occur in parts (i), (ii), (v) and (vi) of 1.3.11 (for all identity morphisms, composable pairs, terminal objects and pullback squares in $\mathcal{C}$). It is easy to see that a structure for $\Sigma_{\mathcal{C}}$ in $\mathcal{D}$ which satisfies the sequents of 1.3.11(i) and (ii) is the same thing as a functor $\mathcal{C} \to \mathcal{D}$, and that it satisfies the sequents of 1.3.11(v) and (vi) iff it preserves the terminal object and pullbacks, i.e. iff it is cartesian.

So, up to equivalence of model categories, cartesian theories are 'the same thing as' small cartesian categories. This is the basis for the notion of Morita equivalence at which we have hinted earlier: we say two cartesian theories $\mathbb{T}$ and $\mathbb{T}'$ are *Morita-equivalent* if their syntactic categories $\mathcal{C}_{\mathbb{T}}$ and $\mathcal{C}_{\mathbb{T}'}$ are equivalent. By 1.4.7, this is equivalent to saying that their categories of models in any cartesian category $\mathcal{D}$ are equivalent, in a fashion which is 'natural in $\mathcal{D}$'. (The term 'Morita equivalence' is borrowed from ring theory, where it is used to describe the relation between two rings which says that their categories of modules are equivalent; the analogy between modules over a ring and models of a theory is often a fruitful one.)

In passing, we note the following result about Morita equivalence, which shows that (for fragments of logic containing cartesian logic) there was some redundancy in Definition 1.1.1, in that we could have done without either the relation symbols or the function symbols.

Lemma 1.4.9 *For any cartesian theory $\mathbb{T}$, there are cartesian theories $\mathbb{T}'$ and $\mathbb{T}''$, Morita-equivalent to $\mathbb{T}$, such that the signature of $\mathbb{T}'$ has no relation symbols, and that of $\mathbb{T}''$ has no function symbols.*

Proof For $\mathbb{T}'$, we could simply take the theory constructed as in 1.4.8 from the syntactic category $\mathcal{C}_{\mathbb{T}}$. Alternatively, we may form a new signature Σ' from that of $\mathbb{T}$, by replacing each relation symbol $R \rightarrowtail A_1 \cdots A_n$ by a new sort $[R]$

together with n function symbols $p_i\colon [R] \to A_i$. Then $\mathbb{T}'$ contains the axiom

$$\Big(\bigwedge_{i=1}^{n}(p_i(w) = p_i(w')) \vdash_{w,w'} (w = w')\Big)$$

(where w and w' are variables of sort $[R]$), for each such R; and whenever a subformula $R(t_1, \dots, t_n)$ appears in an axiom of $\mathbb{T}$, we replace it by the subformula

$$(\exists w) \bigwedge_{i=1}^{n}(p_i(w) = t_i)$$

(note that the sequent displayed above ensures that this formula is cartesian). It is straightforward to verify that every $\mathbb{T}$-model in a cartesian category can be given a Σ'-structure which makes it a $\mathbb{T}'$-model, and vice versa.

For $\mathbb{T}''$, we use the device of replacing functions by the relations which are their graphs: that is, we replace each function symbol $f\colon A_1 \cdots A_n \to B$ in the signature of $\mathbb{T}$ by a new relation symbol $R_f \rightarrowtail A_1 \cdots A_n, B$, for which we adopt the axioms $((R_f(x_1, \dots, x_n, y) \wedge R_f(x_1, \dots, x_n, y')) \vdash (y = y'))$ and $(\top \vdash (\exists y) R_f(x_1, \dots, x_n, y))$ (note that the second of these is cartesian relative to the first). Wherever we have a term of the form $f(x_1, \dots, x_n)$ in a formula ϕ, we may replace the latter by

$$(\exists y)(R_f(x_1, \dots, x_n, y) \wedge \phi')$$

where ϕ' is obtained from ϕ on replacing the term $f(x_1, \dots, x_n)$ by the (new) variable y (note that this is a cartesian formula if ϕ is). Applying this procedure repeatedly, we may eliminate all terms other than variables from the axioms of $\mathbb{T}$; so we arrive at a theory $\mathbb{T}''$, over a signature without function symbols, whose models in any cartesian category are again easily seen to be 'the same as' models of $\mathbb{T}$. □

We remark that cartesian logic is the smallest fragment of first-order logic for which this result is true: the expressive power of Horn logic depends nontrivially on whether we allow our signature to contain function symbols, relation symbols or both – cf. 2.4.4 and 2.4.5 below.

We now consider how the foregoing has to be modified for theories in richer fragments of first-order logic. If $\mathbb{T}$ is a regular (resp. coherent, first-order, geometric) theory, we shall write $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ (resp. $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$) for the category constructed exactly as at the beginning of this section, but allowing the formulae in the definition of both objects and morphisms to range over arbitrary regular (resp. coherent, first-order, geometric) formulae over the relevant signature. The proofs of 1.4.1, 1.4.2 and 1.4.4 remain valid (except for part (ii)

of 1.4.4), and we have

Lemma 1.4.10

(i) *For any regular theory* $\mathbb{T}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ *is a regular category.*

(ii) *For any coherent theory* $\mathbb{T}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ *is a coherent category.*

(iii) *For any first-order theory* $\mathbb{T}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}}$ *is a Heyting category.*

(iv) *For any geometric theory* $\mathbb{T}$, $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ *is a geometric category.*

Proof (i) Given a morphism $[\theta]\colon \{\vec{x}\,.\,\phi\} \to \{\vec{y}\,.\,\psi\}$ in $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$, 1.4.4(iv) implies that we may form its image as the subobject $\{\vec{y}\,.\,(\exists\vec{x})\theta\}$ of $\{\vec{y}\,.\,\psi\}$; in particular, $[\theta]$ is a cover iff the sequent $(\psi \vdash_{\vec{y}} (\exists\vec{x})\theta)$ is provable in $\mathbb{T}$. The fact that covers are stable under pullback follows straightforwardly from the construction of finite limits in the proof of 1.4.2.

(ii) Images in $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ are constructed exactly as in $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$; to form the union of two subobjects $\{\vec{x}\,.\,\psi\}$ and $\{\vec{x}\,.\,\chi\}$ of $\{\vec{x}\,.\,\phi\}$, we of course form $\{\vec{x}\,.\,(\psi \vee \chi)\}$, and the least subobject is $\{\vec{x}\,.\,\bot\}$. It is again easy to verify from 1.4.2 that these constructions are stable under pullback.

(iii) The coherent structure on $\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}}$ is constructed as in (ii). To construct universal quantification along a morphism $[\theta]\colon \{\vec{x}\,.\,\phi\} \to \{\vec{y}\,.\,\psi\}$, let $\{\vec{x}\,.\,\chi\}$ be a subobject of its domain, in the canonical form of 1.4.4(iv). Then we take $\forall_{[\theta]}(\{\vec{x}\,.\,\chi\})$ to be the subobject $\{\vec{y}\,.\,(\psi \wedge (\forall\vec{x})(\theta \Rightarrow \chi))\}$ of $\{\vec{y}\,.\,\psi\}$; it is again straightforward from 1.4.4(iv) to verify that this works.

(iv) is simply the infinitary version of (ii), except that we have to verify that $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ (which is no longer small, because we have a proper class of geometric formulae-in-context) is well-powered. The key to this is the observation made in 1.3.8(ii) that any geometric formula is provably equivalent to a disjunction of regular formulae; since there is only a set of regular formulae in a given context, it follows that there is only a set of provable-equivalence classes of geometric formulae in a given context. This establishes the local smallness of $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ directly from the definition of its morphisms; its well-poweredness follows from 1.4.4(iv). □

Lemma 1.4.5 now extends straightforwardly from cartesian formulae to all formulae in the fragment of the language under consideration, and so we obtain the analogue of 1.4.6:

Theorem 1.4.11 (Completeness Theorem) *Let* $\mathbb{T}$ *be a regular (resp. coherent, first-order, geometric) theory. If* σ *is a regular (resp. coherent, first-order, geometric) sequent which is satisfied in all models of* $\mathbb{T}$ *in regular (resp. coherent, Heyting, geometric) categories, then it is provable in* $\mathbb{T}$. □

The analogue of 1.4.7 also works straightforwardly for regular, coherent and geometric theories. For first-order theories it breaks down, because 1.2.9 fails for non-geometric formulae: thus natural transformations between functors $\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}} \to \mathcal{D}$ correspond not to arbitrary homomorphisms of $\mathbb{T}$-models in $\mathcal{D}$, but to elementary

morphisms as we defined them in 1.2.10(a). (In particular, if $\mathcal{D}$ is Boolean, they correspond to elementary embeddings.)

In passing, we remark that if $\mathbb{T}$ happens to be a cartesian theory, then $\mathcal{C}_{\mathbb{T}}^{\text{reg}}$ is (equivalent to) the regularization $\mathbf{Reg}(\mathcal{C}_{\mathbb{T}})$ of the cartesian syntactic category of $\mathbb{T}$, as constructed in A1.3.9. This follows easily from the fact that regular functors defined on $\mathcal{C}_{\mathbb{T}}^{\text{reg}}$ correspond (to models of $\mathbb{T}$ in the codomain category, and hence) to cartesian functors defined on $\mathcal{C}_{\mathbb{T}}$.

Example 1.4.8 extends in a straightforward way to categories with the richer structure now under consideration. For example, to obtain the theory of regular functors defined on $\mathcal{C}$, we simply add the sequent appearing in 1.3.11(iv), for each cover $f: A \twoheadrightarrow B$ in $\mathcal{C}$, to the theory of 1.4.8, and for the coherent case we add the sequents $(\top \vdash_z \bot)$ where z is a variable of sort 0, and

$$(\top \vdash_x ((\exists y_1)(m_1(y_1) = x) \vee (\exists y_2)(m_2(y_2) = x)))$$

whenever the sort A of x is the union of two subobjects $m_i: A_i \rightarrowtail A$ $(i = 1, 2)$. Lemma 1.4.9 also extends to all the richer fragments of logic.

In general, the category $\mathcal{C}_{\mathbb{T}}^{\text{reg}}$ need not be effective regular, as defined in Section A1.3; nor need $\mathcal{C}_{\mathbb{T}}^{\text{coh}}$ be positive coherent. However, we developed means of effectivizing regular categories, and of positivizing coherent ones, in A3.3.10 and A1.4.5 respectively, and so we may obtain

Proposition 1.4.12

(i) *For any regular theory $\mathbb{T}$, there exists an effective regular category $\mathcal{E}_{\mathbb{T}}$ and a $\mathbb{T}$-model $\overline{M}_{\mathbb{T}}$ in $\mathcal{E}_{\mathbb{T}}$ which is generic among $\mathbb{T}$-models in effective regular categories, in the sense that we have an equivalence*

$$\mathfrak{Reg}\,(\mathcal{E}_{\mathbb{T}}, \mathcal{D}) \simeq \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{D})$$

sending F to $F(\overline{M}_{\mathbb{T}})$, for any effective regular category $\mathcal{D}$. Moreover, the regular sequents satisfied in $\overline{M}_{\mathbb{T}}$ are exactly those provable in $\mathbb{T}$.

(ii) *For any coherent theory $\mathbb{T}$, there is a pretopos $\mathcal{P}_{\mathbb{T}}$ and a $\mathbb{T}$-model $\widetilde{M}_{\mathbb{T}}$ in $\mathcal{P}_{\mathbb{T}}$ which is generic among $\mathbb{T}$-models in pretoposes. Moreover the coherent sequents satisfied in $\widetilde{M}_{\mathbb{T}}$ are exactly those provable in $\mathbb{T}$.*

(iii) *For any geometric theory $\mathbb{T}$, there is an ∞-pretopos $\mathcal{G}_{\mathbb{T}}$ containing a model of $\mathbb{T}$ which is generic among $\mathbb{T}$-models in ∞-pretoposes, and in which satisfaction for geometric sequents is equivalent to provability in $\mathbb{T}$.*

Proof (i) Take $\mathcal{E}_{\mathbb{T}}$ to be the effectivization $\mathbf{Eff}(\mathcal{C}_{\mathbb{T}}^{\text{reg}})$, and $\overline{M}_{\mathbb{T}}$ to be the image of $M_{\mathbb{T}}$ under the canonical embedding. The fact that this embedding is full, and therefore conservative, ensures by 1.2.13 that satisfaction in $\overline{M}_{\mathbb{T}}$ is equivalent to satisfaction in $M_{\mathbb{T}}$; and the generic property of $\overline{M}_{\mathbb{T}}$ follows from combining (the 'regular' version of) 1.4.7 with A3.3.10.

(ii) Similarly, we take $\mathcal{P}_{\mathbb{T}}$ to be $\mathbf{Eff}(\mathbf{Pos}(\mathcal{C}_{\mathbb{T}}^{\text{coh}}))$, where $\mathbf{Pos}$ is the positivization functor defined in A1.4.5. Once again, the canonical coherent functor

$\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}} \to \mathcal{P}_{\mathbb{T}}$ is full and faithful, and therefore conservative; the remaining details are similar to (i).

(iii) is similar to (ii), except that we use the infinitary analogue of the positivization construction, mentioned at the end of Section A1.4. Note that this construction, and the effectivization construction, do not destroy well-poweredness. □

Since the embeddings of the syntactic categories into the categories considered in 1.4.12 are full and therefore conservative, it follows from 1.2.13 that we have a completeness theorem for regular (resp. coherent, geometric) logic involving only models in effective regular categories (resp. pretoposes, ∞-pretoposes).

We define *Morita equivalence* for regular (resp. coherent, geometric) theories to mean equivalence of the categories $\mathcal{E}_{\mathbb{T}}$ (resp. $\mathcal{P}_{\mathbb{T}}$, $\mathcal{G}_{\mathbb{T}}$), rather than of the syntactic categories themselves. Of course, this is a weaker notion than equivalence of the syntactic categories; but it suffices to ensure that the theories have equivalent model categories in any (cocomplete) topos, which is our main concern. (We shall see in 3.1.12 and 3.3.8 below that the converse is true.)

Next we note a further simplification of the notion of signature, analogous to 1.4.9, which may be made in the context of geometric logic.

Lemma 1.4.13 *Let $\mathbb{T}$ be a geometric theory. Then there is a geometric theory $\mathbb{T}'$, Morita-equivalent to $\mathbb{T}$, whose signature has only one sort.*

Proof First we assume, by using 1.4.9 if necessary, that the signature Σ of $\mathbb{T}$ has no function symbols. Let $(A_i \mid i \in I)$ be the sorts of Σ; then our new signature has a single sort A and a family of unary relation symbols $U_i \rightarrowtail A$ corresponding to the sorts of Σ, together with a relation symbol $\overline{R} \rightarrowtail A \cdots A$ for each relation symbol $R \rightarrowtail A_1 \cdots A_n$ of Σ. Among the axioms of $\mathbb{T}'$, we have $((U_i(x) \wedge U_j(x)) \vdash_x \bot)$ whenever $i \neq j$, and $(\top \vdash_x \bigvee_{i \in I} U_i(x))$, which say that the interpretation $[\![x \,.\, \top]\!]$ in a $\mathbb{T}'$-model is the disjoint coproduct of the interpretations $[\![x \,.\, U_i(x)]\!]$. Then, for each $R \rightarrowtail A_1 \cdots A_n$ in Σ, we adopt the axiom

$$\Big(\overline{R}(x_1, \ldots, x_n) \vdash_{x_1, \ldots, x_n} \bigwedge_{i=1}^{n} U_i(x_i)\Big),$$

and finally each axiom of $\mathbb{T}$ gives rise to an axiom of $\mathbb{T}'$ by putting lines over all the relation symbols, and introducing, on the left of each sequent, the additional hypotheses that the variables in its context satisfy the new unary relations corresponding to the sorts which they had in Σ. It is then easy to see that a $\mathbb{T}$-model M in an ∞-pretopos $\mathcal{G}$ gives rise to a $\mathbb{T}'$-model whose underlying object is the coproduct of the MA_i, and conversely; so the ∞-pretoposes $\mathcal{G}_{\mathbb{T}}$ and $\mathcal{G}_{\mathbb{T}'}$ must be equivalent. □

We remark that, in contrast to 1.4.9, geometric logic is not the smallest fragment of infinitary first-order logic for which this reduction is possible; we could have done the same thing in the disjunctive logic mentioned in 1.3.6.

Remark 1.4.14 Before leaving this section, we should briefly discuss the syntactic categories of propositional theories as defined in 1.1.7(m). Since a propositional theory $\mathbb{T}$ has only one (empty) context, it is easy to see that $\mathcal{C}_{\mathbb{T}}$ is a preorder: there exists a morphism $\{[\,].\phi\} \to \{[\,].\psi\}$ iff $(\phi \vdash_{[]} \psi)$ is provable in $\mathbb{T}$, and the only such morphism is then $[\phi]$. Thus $\mathcal{C}_{\mathbb{T}}$ is equivalent to what is commonly called the *Lindenbaum algebra* of $\mathbb{T}$, that is the poset of provable-equivalence classes of formulae, ordered by provable entailment. (For a general $\mathbb{T}$, the Lindenbaum algebra – defined as above, but restricted to formulae in the empty context, i.e. to sentences – appears as the poset of subterminal objects in $\mathcal{C}_{\mathbb{T}}$.) The structure carried by the Lindenbaum algebra of course depends on the level of logical complexity at which $\mathbb{T}$ is considered to live: if $\mathbb{T}$ is a cartesian theory then it is a meet-semilattice, if $\mathbb{T}$ is coherent then it is a distributive lattice, if $\mathbb{T}$ is first-order then it is a Heyting algebra, and if $\mathbb{T}$ is geometric then it is a frame (the order-theoretic operations in each case being induced by the corresponding logical operations on formulae). Moreover, it is easily seen that if L is a meet-semilattice (resp. a distributive lattice, a frame) and $\mathcal{C}$ is an arbitrary cartesian (resp. coherent, geometric) category, then cartesian (resp. coherent, geometric) functors $L \to \mathcal{C}$ correspond bijectively to models in $\mathcal{C}$ of the theory of filters (resp. prime filters, completely prime filters) of L, as defined in 1.1.7(m) (note that, since such functors preserve monomorphisms, they all take values in the poset of subterminal objects of $\mathcal{C}$). Thus we may conclude that *every cartesian (resp. coherent, geometric) propositional theory is Morita-equivalent to the theory of filters (resp. prime filters, completely prime filters) of a meet-semilattice (resp. a distributive lattice, a frame).*

Remark 1.4.15 The sequents appearing in the definition of morphisms of $\mathcal{C}_{\mathbb{T}}$, earlier in this section, may have reminded readers of the construction of the topos **Set**(L) from the allegory **Mat**(L) in Section A3.3. This is no accident: for fragments of logic containing regular logic, the syntactic category may conveniently be constructed by way of a 'syntactic allegory', which is in some ways a more natural generalization of the Lindenbaum algebra mentioned in the preceding remark than the syntactic category itself. Formally, if $\mathbb{T}$ is (say) a regular theory, we take $\mathcal{A}_{\mathbb{T}}^{\mathrm{reg}}$ to be the allegory whose objects are α-equivalence classes of *contexts* over the signature of $\mathbb{T}$, and whose morphisms from $\vec{x}$ to $\vec{y}$ (where these contexts are presumed disjoint) are $\mathbb{T}$-provable equivalence classes of regular formulae in the concatenated context $\vec{x}, \vec{y}$. These morphisms are ordered by $\mathbb{T}$-provable entailment (i.e. $[\phi] \leq [\psi]$ iff $(\phi \vdash_{\vec{x},\vec{y}} \psi)$ is provable in $\mathbb{T}$); the opposite of $[\phi]$ is simply $[\phi]$ in the re-ordered context $\vec{y}, \vec{x}$; and the composite of $[\phi]\colon \vec{x} \nrightarrow \vec{y}$ and $[\psi]\colon \vec{y} \nrightarrow \vec{z}$ is $[(\exists \vec{y})(\phi \wedge \psi)]$. It is readily verified that this is an allegory, and that it has a unit (the empty context) and is pre-tabular (the top element of $\mathcal{A}_{\mathbb{T}}^{\mathrm{reg}}(\vec{x}, \vec{y})$ is $[\top]$, and it is tabulated by

$$\vec{x} \xleftarrow{[(\vec{x} = \vec{x'})]} \vec{x'}, \vec{y'} \xrightarrow{[(\vec{y} = \vec{y'})]} \vec{y}) \,.$$

So by A3.3.6(ii) and A3.2.10 $\mathbf{Map}(\mathcal{A}_{\mathbb{T}}^{\mathrm{reg}}[\check{\mathcal{K}}])$ is a regular category, which is simply $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ as we have defined it in this section (and $\mathbf{Map}(\mathcal{A}_{\mathbb{T}}^{\mathrm{reg}}[\check{\mathcal{S}}])$ is its effectivization $\mathcal{E}_{\mathbb{T}}$). Similarly, for a coherent (resp. first-order, geometric) theory $\mathbb{T}$ we obtain a union (resp. division, geometric) allegory $\mathcal{A}_{\mathbb{T}}^{\mathrm{coh}}$ (resp. $\mathcal{A}_{\mathbb{T}}^{\mathrm{fo}}$, $\mathcal{A}_{\mathbb{T}}^{\mathrm{geom}}$), which can be used to construct the appropriate syntactic category in each case. Unfortunately, however, this method will not work for cartesian theories, since by A3.1.1 we cannot hope to obtain a non-regular cartesian category as the category of maps of an allegory.

Suggestions for further reading: Freyd et al. [377], Makkai & Reyes [790].

D1.5 Classical completeness

In the previous section we proved completeness theorems (1.4.6 and 1.4.11) in the form: for any theory $\mathbb{T}$, there is a single model $M_{\mathbb{T}}$ such that satisfaction in $M_{\mathbb{T}}$ is equivalent to provability in $\mathbb{T}$. Despite the evident neatness of this form of the theorem, the fact that $M_{\mathbb{T}}$ lives (not in **Set** but) in a rather 'exotic' category, constructed specially for the purpose of proving the theorem, is sometimes seen as a disadvantage. In 'classical' accounts of first-order logic, the completeness theorem involves only models in the 'pre-existing' category **Set**; so it is of interest to see whether such a theorem can be derived from the one we have proved.

For cartesian theories, the answer is an immediate 'yes': we know from A1.2.2 that for any small cartesian category $\mathcal{C}$ there is a jointly-conservative family of cartesian functors $\mathcal{C} \to \mathbf{Set}$, namely the representable functors $\mathcal{C}(A, -)$, $A \in \mathrm{ob}\, \mathcal{C}$. Applying this to $\mathcal{C}_{\mathbb{T}}$, and appealing to 1.2.13, we immediately obtain

Proposition 1.5.1 *Let $\mathbb{T}$ be a cartesian theory, and σ a cartesian sequent relative to $\mathbb{T}$. If σ is satisfied in all $\mathbb{T}$-models in* **Set***, then it is provable in $\mathbb{T}$.*

Proof If σ is satisfied in the $\mathbb{T}$-models which correspond via 1.4.7 to the representable functors, then it is satisfied in $M_{\mathbb{T}}$ by 1.2.13, and hence provable in $\mathbb{T}$ by 1.4.6. □

Our strategy for proving the classical completeness theorems for regular and coherent logic will be similar to the above: we shall show that for every small regular (resp. coherent) category $\mathcal{C}$, there is a (set-indexed) family of regular (resp. coherent) functors $\mathcal{C} \to \mathbf{Set}$ which is jointly conservative. There are various ways of achieving this: we shall follow an approach due to P. Freyd (see [381]), which has the advantage of minimizing the amount of duplication of effort required for the two cases.

We recall the notion of capital regular category introduced in A1.3.8: our first aim is to show that every small regular category $\mathcal{C}$ admits a conservative regular functor to a (small) capital regular category $\check{\mathcal{C}}$.

We define a *base* to be a finite sequence $\vec{B} = (B_1, B_2, \ldots, B_n)$ of *distinct* well-supported objects of $\mathcal{C}$. (The requirement that the B_i be distinct seems rather

bizarre, given that it doesn't matter how many isomorphic copies of any given object there are in $\mathcal{C}$; but it appears to be essential for technical reasons.) Given a base $\vec{B}$, we write $\prod \vec{B}$ for the (canonical) product $B_1 \times \cdots \times B_n$ in $\mathcal{C}$; the need to have a well-defined product is the only reason for ordering the members of a base. We preorder the set of bases by setting $\vec{B} \leq \vec{C}$ if every object in $\vec{B}$ appears in $\vec{C}$ (but not necessarily in the same order); clearly, the preordered set of bases is directed. If $\vec{B} \leq \vec{C}$ and we are given an object $f: A \to \prod \vec{B}$ of $\mathcal{C}/\prod \vec{B}$, we write $\vec{C}^*A$ for the product $A \times \prod C_j$ (where j runs over those indices such that C_j is *not* in $\vec{B}$), i.e. the pullback of A along the appropriate product projection $\prod \vec{C} \to \prod \vec{B}$, and we write $\vec{C}^*f : \vec{C}^*A \to \prod \vec{C}$ for the pullback of f along this projection.

We now define a category $\mathcal{C}_1$ as follows. Its objects are triples $(A, \vec{B}, f)$ where A is an object of $\mathcal{C}$, $\vec{B}$ is a base and $f: A \to \prod \vec{B}$. Morphisms $(A, \vec{B}, f) \to (A', \vec{B}', f')$ are represented by pairs $(\vec{C}, g)$ where $\vec{C}$ is a base containing both $\vec{B}$ and $\vec{B}'$ and $g: \vec{C}^*f \to \vec{C}^*f'$ in $\mathcal{C}/\prod \vec{C}$, modulo the equivalence relation which identifies $(\vec{C}, g)$ and $(\vec{C}', g')$ if there is a base $\vec{C}''$ containing both $\vec{C}$ and $\vec{C}'$, such that g and g' pull back to the same morphism in $\mathcal{C}/\prod \vec{C}''$. (Recall that pullback along a cover is conservative by A1.3.2(iii); so the pairs $(\vec{C}, g)$ and $(\vec{C}, g')$ represent the same morphism iff $g = g'$, and more importantly $(\vec{C}, g)$ is an isomorphism in $\mathcal{C}_1$ iff g is an isomorphism in $\mathcal{C}/\prod \vec{C}$.) Composition and identity morphisms are defined in the obvious way; it is clear that $\mathcal{C}_1$ is a category.

If $\vec{B} \leq \vec{C}$, then the object $(A, \vec{B}, f)$ is isomorphic to $(\vec{C}^*A, \vec{C}, \vec{C}^*f)$ in $\mathcal{C}_1$ (the isomorphism being represented by $(\vec{C}, 1_{\vec{C}^*A})$ in each direction). Hence if, for a fixed $\vec{B}$, we consider the (non-full) subcategory $\mathcal{C}_1|\vec{B}$ of $\mathcal{C}_1$ whose objects are those $(A, \vec{B}', f)$ with $\vec{B}' \leq \vec{B}$, and whose morphisms are representable by pairs of the form $(\vec{B}, g)$, we see that $\mathcal{C}_1|\vec{B}$ is equivalent to $\mathcal{C}/\prod \vec{B}$ (the equivalence sends $(A, \vec{B}', f)$ to $\vec{B}^*f$ in one direction, and f to $(\operatorname{dom} f, \vec{B}, f)$ in the other). In particular, $\mathcal{C}_1|[\,]$ (where $[\,]$ denotes the empty base) is equivalent to $\mathcal{C}$ itself. Moreover, if $\vec{B} \leq \vec{C}$, then this equivalence identifies the inclusion $\mathcal{C}_1|\vec{B} \subseteq \mathcal{C}_1|\vec{C}$ with the pullback functor $\vec{C}^* : \mathcal{C}/\prod \vec{B} \to \mathcal{C}/\prod \vec{C}$.

But $\mathcal{C}_1$ is clearly the union of the subcategories $\mathcal{C}_1|\vec{B}$, as $\vec{B}$ ranges over all bases; so we have expressed it as a directed union of regular subcategories and conservative regular functors between them. Thus we obtain

Proposition 1.5.2 *For any small regular category $\mathcal{C}$, there exists a small regular category $\mathcal{C}_1$ and a conservative regular functor $T: \mathcal{C} \to \mathcal{C}_1$, such that if $A' \rightarrowtail A$ is any proper subobject in $\mathcal{C}$ with A well-supported, there exists a morphism $1 \to TA$ in $\mathcal{C}_1$ not factoring through $TA' \rightarrowtail TA$. Moreover, if $\mathcal{C}$ is (positive) coherent, then so is $\mathcal{C}_1$ and the functor T is coherent.*

Proof The construction of $\mathcal{C}_1$, and the proof that it is regular (and positive coherent if $\mathcal{C}$ is), is contained in the discussion above. The functor T is obtained

by identifying $\mathcal{C}$ with $\mathcal{C}_1|[\,] \subseteq \mathcal{C}_1$; the fact that it is conservative follows from the fact that the functors $\mathcal{C} \to \mathcal{C}/\prod \vec{B}$ are all conservative, since if Tf has an inverse in $\mathcal{C}_1$ then the latter must lie in $\mathcal{C}_1|\vec{B}$ for some $\vec{B}$. Finally, given $A' \rightarrowtail A$ as in the statement of the proposition, we note that the singleton sequence (A) is a base, and that in $\mathcal{C}_1|(A) \simeq \mathcal{C}/A$ we already have a morphism $1 \to TA$ not factoring through TA', namely the diagonal $A \rightarrowtail A \times A$ regarded as a morphism $1_A \to A^*A$ in $\mathcal{C}/A$. □

The category $\mathcal{C}_1$ constructed in 1.5.2 will not in general be capital; for its well-supported objects will have proper subobjects which are not in the image of the functor T. However, we can achieve the capital property by iterating the construction just described:

Theorem 1.5.3 *For any small regular category $\mathcal{C}$, there exists a small capital regular category $\widehat{\mathcal{C}}$ and a conservative regular functor $F\colon \mathcal{C} \to \widehat{\mathcal{C}}$. Moreover, if $\mathcal{C}$ is (positive) coherent, then so is $\widehat{\mathcal{C}}$ and the functor F is coherent.*

Proof We construct a sequence of categories and functors

$$\mathcal{C}_0 \xrightarrow{T_0} \mathcal{C}_1 \xrightarrow{T_1} \mathcal{C}_2 \xrightarrow{T_2} \cdots$$

where $\mathcal{C}_0 = \mathcal{C}$ and $\mathcal{C}_{n+1}$ is constructed from $\mathcal{C}_n$ by the method of 1.5.2. We take $\widehat{\mathcal{C}}$ to be the (pseudo-)colimit of this sequence in $\mathfrak{Cat}$: explicitly, objects of $\widehat{\mathcal{C}}$ are pairs (n, A) with $A \in \text{ob}\,\mathcal{C}_n$, and morphisms $(n, A) \to (n', A')$ are represented by pairs (m, f) where $m \geq \max\{n, n'\}$ and g is a morphism between the images of A and A' in $\mathcal{C}_m$, modulo the obvious equivalence relation. As before, the fact that the diagram is filtered ensures that its colimit inherits any finitary categorical structure (such as regularity or coherence) enjoyed by all the categories $\mathcal{C}_n$ and preserved by the transition functors T_n; and since all these functors are conservative, so are the induced functors $\mathcal{C}_n \to \widehat{\mathcal{C}}$ for all n – in particular for $n = 0$. Finally, $\widehat{\mathcal{C}}$ is capital because, if $f\colon A' \rightarrowtail A$ is any proper monomorphism in it with A well-supported, there exists n such that f lies in the image of the functor $\mathcal{C}_n \to \widehat{\mathcal{C}}$; and then in $\mathcal{C}_{n+1}$ we can find a morphism $1 \to A$ which does not factor through f. □

Corollary 1.5.4

(i) *For any small regular category $\mathcal{C}$, there is a conservative regular functor $\mathcal{C} \to$ **Set**$/B$ for some set B.*

(ii) **(Classical completeness for regular logic)** *If $\mathbb{T}$ is a regular theory and σ is a regular sequent over the signature of $\mathbb{T}$ which is satisfied in all $\mathbb{T}$-models in* **Set**, *then σ is provable in $\mathbb{T}$.*

Proof (i) Consider the family of composite functors

$$\mathcal{C} \xrightarrow{U^*} \mathcal{C}/U \xrightarrow{F} \widehat{\mathcal{C}/U} \xrightarrow{\Gamma} \mathbf{Set}$$

where U ranges over the set of subterminal objects of $\mathcal{C}$, F is the 'capitalization' of $\mathcal{C}/U$ constructed in 1.5.3, and Γ is the functor $\widehat{\mathcal{C}/U}(1, -)$. By A1.3.8, the terminal object of $\widehat{\mathcal{C}/U}$ is projective, so Γ (and hence also the composite above) is a regular functor. Moreover, if $A' \rightarrowtail A$ is a proper subobject in $\mathcal{C}$, let $U \rightarrowtail 1$ denote the support of A; then $U^*A' \rightarrowtail U^*A$ is still proper, and U^*A is well-supported in $\mathcal{C}/U$, so (for this particular U) the composite above maps $A' \rightarrowtail A$ to a proper monomorphism. Hence by A1.2.4 the family of all such composites is jointly conservative; equivalently, they can be combined into a single conservative functor $\mathcal{C} \to \mathbf{Set}/B$ where B is a set indexing the subterminal objects of $\mathcal{C}$.

(ii) We apply (i) to the syntactic category $\mathcal{C}_{\mathbb{T}}^{\text{reg}}$ of $\mathbb{T}$. If σ is satisfied in the **Set**-models of $\mathbb{T}$ which correspond to the functors in the conservative family constructed in (i), then by 1.2.13 it is satisfied in the generic model $M_{\mathbb{T}}$, and hence provable in $\mathbb{T}$. □

Remark 1.5.5 As we mentioned after A1.3.8, there is a significant strengthening of the first part of 1.5.4: every small regular category $\mathcal{C}$ admits a full and faithful regular functor into a functor category $[\mathcal{D}, \mathbf{Set}]$, where $\mathcal{D}$ is a small category. However, the extra strength of this result does not yield any significant extra information about regular theories, and so we shall not prove it here.

Remark 1.5.6 It is of interest to ask when a (small) regular category $\mathcal{C}$ admits a single conservative regular functor to **Set** (i.e. when we can take $B = 1$ in 1.5.4(i)). This will clearly be the case if 1 is the only subterminal object of $\mathcal{C}$, i.e. if every object of $\mathcal{C}$ is well-supported; but it will also hold if (as in **Set** itself) the only proper subobject of 1 is a strict initial object 0, since then $\mathcal{C}/0$ is degenerate and we can omit the functor corresponding to 0 from the family of 1.5.4(i) without affecting its conservativity. So we arrive at the sufficient condition that *given any diagram of the form*

$$A \xrightarrow{f} B \overset{g}{\rightarrowtail} 1$$

in $\mathcal{C}$, either f or g must be an isomorphism. But this condition holds in **Set**, and its truth is clearly reflected by conservative regular functors; so the condition is necessary as well as sufficient.

For coherent categories, we have to do a bit more work; for the functors Γ used in the proof of 1.5.4(i), though regular, will not in general be coherent. The key to obtaining coherent **Set**-valued functors is the following result:

Lemma 1.5.7 *In a positive coherent category, a complemented subobject of a (cover-)projective object is projective.*

Proof Let $A' \rightarrowtail A$ be a complemented subobject of a projective object A, with complement $A'' \rightarrowtail A$. We have to show that every cover $e\colon B \twoheadrightarrow A'$ splits; but any such cover extends to a cover $e \amalg 1\colon B \amalg A'' \twoheadrightarrow A' \amalg A'' \cong A$, and by disjointness of coproducts a splitting for this must map A' into B, so it yields a splitting for e. □

Thus, in a capital positive coherent category $\mathcal{C}$, every complemented subterminal object U is projective – equivalently, the functor $\mathcal{C}(U, -)$ is regular. Given such a category, let $\mathcal{B}$ denote the lattice (in fact a Boolean algebra) of complemented subterminal objects of $\mathcal{C}$; and for any filter $\mathcal{F} \subseteq \mathcal{B}$, let $\Gamma_{\mathcal{F}}$ denote the (filtered) colimit (in $[\mathcal{C}, \mathbf{Set}]$) of the functors $\mathcal{C}(U, -)$, $U \in \mathcal{F}$. Explicitly, $\Gamma_{\mathcal{F}}(A)$ is the quotient of $\coprod_{U \in \mathcal{F}} \mathcal{C}(U, A)$ by the equivalence relation which identifies $f\colon U \to A$ with $g\colon V \to A$ if there exists a commutative square

$$\begin{tikzcd} W \arrow[r, tail] \arrow[d, tail] & U \arrow[d, "f"] \\ V \arrow[r, "g"] & A \end{tikzcd}$$

with $W \in \mathcal{F}$. It is clear that $\Gamma_{\mathcal{F}}$, as a filtered colimit of regular functors, is again regular.

Lemma 1.5.8 *The functor $\Gamma_{\mathcal{F}}$ constructed above is coherent iff $\mathcal{F}$ is an ultrafilter (i.e. a prime filter) in $\mathcal{B}$.*

Proof Since $\mathcal{C}$ is positive coherent, $\Gamma_{\mathcal{F}}$ will be coherent iff it preserves finite coproducts. But $\Gamma_{\mathcal{F}}$ preserves 0 iff $0 \notin \mathcal{F}$; and, given an element x of $\Gamma_{\mathcal{F}}(A \amalg B)$ (represented by a morphism $f\colon U \to A \amalg B$, say), we can decompose U as a disjoint union $f^*A \amalg f^*B$, and primeness says that just one of these will be in $\mathcal{F}$, so x is in (the image of) just one of $\Gamma_{\mathcal{F}}(A)$ and $\Gamma_{\mathcal{F}}(B)$. Thus $\Gamma_{\mathcal{F}}$ preserves binary coproducts if $\mathcal{F}$ is prime; for the converse, note that a complemented subterminal object U satisfies $\Gamma_{\mathcal{F}}(U) \cong 1$ iff $U \in \mathcal{F}$. □

Proposition 1.5.9 *Let $\mathcal{C}$ be a small capital positive coherent category. Then the functors $\Gamma_{\mathcal{F}}$, as $\mathcal{F}$ ranges over all prime filters of complemented subterminal objects of $\mathcal{C}$, are jointly conservative.*

Proof Let $A' \rightarrowtail A$ be a proper subobject in $\mathcal{C}$. Although A may not be well-supported, $A \amalg 1$ is, and so we can find $f\colon 1 \to A \amalg 1$ not factoring through $A' \amalg 1 \rightarrowtail A \amalg 1$. Pulling back along the first coprojection, we obtain a complemented subterminal object U and $g\colon U \to A$ not factoring through A'. Form the

further pullback

$$\begin{tikzcd} V \arrow[r] \arrow[d, tail] & A' \arrow[d, tail] \\ U \arrow[r, "g"] & A \end{tikzcd}$$

and let $\mathcal{I}$ be the ideal of complemented subterminal objects contained in V. (Note that V itself need not be complemented.) Since $U \notin \mathcal{I}$, the Boolean Prime Ideal Theorem tells us that there is a prime filter $\mathcal{F}$ with $U \in \mathcal{F}$ and $\mathcal{F} \cap \mathcal{I} = \emptyset$; then g represents an element of $\Gamma_{\mathcal{F}}(A)$ which does not belong to $\Gamma_{\mathcal{F}}(A')$. □

Corollary 1.5.10

(i) *For any small coherent category $\mathcal{C}$, there is a conservative coherent functor $\mathcal{C} \to \mathbf{Set}/B$ for some set B.*

(ii) **(Classical completeness for coherent logic)** *If $\mathbb{T}$ is a coherent theory, and σ is a coherent sequent over the signature of $\mathbb{T}$ which is satisfied in all $\mathbb{T}$-models in* **Set**, *then σ is provable in $\mathbb{T}$.*

Proof (i) Consider the composite functors

$$\mathcal{C} \longrightarrow \mathbf{Pos}(\mathcal{C}) \xrightarrow{F} \widehat{\mathbf{Pos}(\mathcal{C})} \xrightarrow{\Gamma_{\mathcal{F}}} \mathbf{Set}$$

where $\mathbf{Pos}(\mathcal{C})$ is the positivization of $\mathcal{C}$ constructed in A1.4.5, F is its capitalization as constructed in 1.5.3, and $\mathcal{F}$ ranges over all prime filters of complemented subterminal objects in $\widehat{\mathbf{Pos}(\mathcal{C})}$. By 1.5.3, $\widehat{\mathbf{Pos}(\mathcal{C})}$ is positive coherent as well as capital, so by 1.5.8 and 1.5.9 the functors $\Gamma_{\mathcal{F}}$ (and hence also the composites above) are coherent and jointly conservative. (Recall that the embedding $\mathcal{C} \to \mathbf{Pos}(\mathcal{C})$ is full and hence conservative.) So we may combine them, as before, into a single conservative coherent functor $\mathcal{C} \to \mathbf{Set}/B$, where B is a set indexing the filters $\mathcal{F}$.

(ii) follows from (i) exactly as 1.5.4(ii) followed from 1.5.4(i). □

Remark 1.5.11 Before proceeding further, we should make some comments about the use of the axiom of choice in the proofs of 1.5.4 and 1.5.10. Although the construction of 1.5.2 does not make any explicit use of choice, it does require the existence of a canonical choice of products in the category $\mathcal{C}$ from which we start. In the category $\mathcal{C}_1$ constructed in 1.5.2, products are constructed using pullbacks in $\mathcal{C}$, so we will have a canonical choice of them provided $\mathcal{C}$ has a canonical choice of pullbacks; however, the construction of pullbacks in $\mathcal{C}_1$ requires the making of arbitrary choices, because of the equivalence relation involved in the definition of morphisms of $\mathcal{C}_1$. Thus the iteration of the construction of 1.5.2, which is performed in the proof of 1.5.3, will eventually force us to make arbitrary choices, unless we assume at the outset that we are given something like

a well-ordering of the objects and morphisms of $\mathcal{C}$ – it is not hard to see that such a well-ordering can be 'carried along' through all the stages of the iteration. Of course, if we start from the category $\mathcal{C}_{\mathbb{T}}^{\text{reg}}$, where $\mathbb{T}$ is a regular theory over a countable signature Σ, it is not hard to construct such a well-ordering; so we may deduce that the classical completeness theorem for regular theories over countable signatures is valid without any use of choice.

On the other hand, for coherent logic the use of the Boolean Prime Ideal Theorem, to which we appealed explicitly in the proof of 1.5.9, is unavoidable. In fact the classical completeness theorem for coherent logic is equivalent to this theorem, as may be seen by applying it to the propositional theory of prime filters of a distributive lattice L, which we described in 1.1.7(m). Since this theory has a model (in fact its generic model, up to equivalence) in L itself considered as a coherent category, it is consistent whenever L is non-degenerate (i.e. satisfies $0 \neq 1$), and hence classical completeness implies that every non-degenerate distributive lattice has a prime filter – which is well known to be one of the equivalent forms of the Prime Ideal Theorem. (See [520] for a discussion of the various forms in which this theorem appears in algebra and topology.) To put the same thing in more categorical terms: the assertion, for a distributive lattice L considered as a coherent category, that the coherent functors $L \to \mathbf{Set}$ are jointly conservative is equivalent to saying that the prime filters of L separate its elements – which is another form of the theorem.

Remark 1.5.12 By an argument similar to that just given, we may see that there can be no analogue of 1.5.10(ii) for geometric theories. For if we take $\mathbb{T}$ to be a geometric propositional theory, then $\mathbb{T}$ has a model in its Lindenbaum algebra L (cf. 1.4.14), or equivalently in the topos $\mathbf{Sh}(X)$, where X is the locale corresponding to L (cf. 1.2.15(m)), and so it is consistent whenever L is non-degenerate. But its models in **Set** are the same thing as points of X; so if X is a nontrivial locale without points (for example, if $\mathbb{T}$ is either of the theories of C1.2.8 or C1.2.9), then $\mathbb{T}$ is a consistent geometric theory having no models in **Set**.

Nevertheless, there is a 'classical completeness theorem' for geometric logic, which asserts that any geometric theory has enough models to determine provability (not in **Set** but) in Boolean toposes. We shall prove this theorem in 3.1.16 below.

There can be no completeness theorem in terms of **Set**-models (or even in terms of models in Boolean toposes) for (constructive) full first-order logic; for we know that the Law of Excluded Middle (1.3.3) is satisfied in all such models, but not provable in this logic. However, if we add this law as an axiom, thus obtaining classical first-order logic, we do have a completeness theorem, which we can deduce from that for coherent logic by the following trick.

Lemma 1.5.13 *Let $\mathbb{T}$ be a first-order theory over a signature Σ. Then there is a signature Σ' containing Σ, and a coherent theory $\mathbb{T}'$ over Σ', such that for*

any Boolean coherent category $\mathcal{C}$ we have

$$\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{C})_e \simeq \mathbb{T}'\text{-}\mathbf{Mod}(\mathcal{C}) \, .$$

The theory $\mathbb{T}'$ is sometimes called the *Morleyization* of $\mathbb{T}$, in honour of M. Morley (cf. [446]).

Proof We define Σ' to have the same sorts and function symbols as Σ, but (in addition to the relation symbols of Σ) we add two new relation symbols $C_\phi \rightarrowtail A_1 \cdots A_n$ and $D_\phi \rightarrowtail A_1 \cdots A_n$ for each first-order formula ϕ over Σ (here the type $A_1 \cdots A_n$ of C_ϕ and D_ϕ is the string of sorts corresponding to the canonical context of ϕ; we could have introduced new relation symbols for all the formulae-in-context over Σ, but thanks to the Substitution Lemma 1.2.7 there is no need to do so). In $\mathbb{T}'$ we include the axioms

$$((C_\phi(x_1, \ldots, x_n) \wedge D_\phi(x_1, \ldots, x_n)) \vdash_{x_1, \ldots, x_n} \bot)$$

and

$$(\top \vdash_{x_1, \ldots, x_n} (C_\phi(x_1, \ldots, x_n) \vee D_\phi(x_1, \ldots, x_n)))$$

which ensure that the interpretations of C_ϕ and D_ϕ in any $\mathbb{T}'$-model M are complementary subobjects of $M(A_1, \ldots, A_n)$, for each formula ϕ over Σ. Then we have a group of axioms, depending on the structure of the formula ϕ, which specify the interpretation of the relation symbol C_ϕ in terms of the C_ψ for appropriate subformulae ψ of ϕ, as follows:

(i) If ϕ is atomic,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} \phi)$$

(where we use the notation introduced in 1.3.7 to indicate that we are taking both a sequent and its converse as axioms);

(ii) If $\phi = \top$,

$$(\top \vdash_{[]} C_\phi);$$

(iii) If $\phi = (\psi \wedge \chi)$,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} (C_\psi \wedge C_\chi))$$

(here the relation symbols on the right are applied to the appropriate substrings of $\vec{x}$);

(iv) If $\phi = \bot$,

$$(C_\phi \vdash_{[]} \bot);$$

(v) If $\phi = (\psi \vee \chi)$,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} (C_\psi \vee C_\chi));$$

(vi) If $\phi = \neg\psi$,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} D_\psi(\vec{x}));$$

(vii) If $\phi = (\psi \Rightarrow \chi)$,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} (D_\psi \vee C_\chi));$$

(viii) If $\phi = (\exists y)\psi$,

$$(C_\phi(\vec{x}) \dashv\vdash_{\vec{x}} (\exists y)C_\psi(\vec{x}, y));$$

(ix) If $\phi = (\forall y)\psi$,

$$(D_\phi(\vec{x}) \dashv\vdash_{\vec{x}} (\exists y)D_\psi(\vec{x}, y)) .$$

It will be seen that these axioms are all coherent, and that they ensure, for any $\mathbb{T}'$-model M in a Boolean coherent category, that the interpretation of $\vec{x}.C_\phi(\vec{x})$ coincides with that of $\vec{x}.\phi$ (for any suitable context $\vec{x}$, not just the canonical context of ϕ). Thus, to ensure that $\mathbb{T}'$-models coincide with $\mathbb{T}$-models in such categories, we have simply to add the axiom $(C_\phi \vdash_{\vec{x}} C_\psi)$ for each axiom $(\phi \vdash_{\vec{x}} \psi)$ of $\mathbb{T}$. The fact that homomorphisms of $\mathbb{T}'$-models correspond to elementary morphisms of $\mathbb{T}$-models follows from the fact that the former must preserve the interpretations of all the C_ϕ. □

Corollary 1.5.14 (Classical completeness for first-order logic) *Let $\mathbb{T}$ be a first-order theory over a signature Σ. If a first-order sequent $(\phi \vdash_{\vec{x}} \psi)$ over Σ is satisfied in all $\mathbb{T}$-models in* **Set**, *then it is provable in $\mathbb{T}$ using classical logic.*

Proof If $(\phi \vdash_{\vec{x}} \psi)$ is satisfied in all $\mathbb{T}$-models in **Set**, then (in the notation of 1.5.13) $(C_\phi \vdash_{\vec{x}} C_\psi)$ is satisfied in all $\mathbb{T}'$-models in **Set**, and so is provable in $\mathbb{T}'$. But if we work through this proof replacing every subformula of the form $C_\chi(t_1, \ldots, t_m)$ by the corresponding substitution instance $\chi[\vec{t}/\vec{x}]$ of χ, and every D_χ by a substitution instance of $\neg\chi$, we will obtain a classical proof in $\mathbb{T}$ of the sequent $(\phi \vdash \psi)$, since the effect of this transformation on each of the axioms of $\mathbb{T}'$ is to produce either a classically derivable sequent or an axiom of $\mathbb{T}$. For example, the second axiom in the first group becomes the Law of Excluded Middle itself; and the left-to-right sequent (ix) of the second group transforms into $(\neg(\forall y)\psi \vdash_{\vec{x}} (\exists y)\neg\psi)$, which may be proved by combining the following (constructively valid) derivation with the instance $(\top \vdash_{\vec{x}} ((\exists y)\neg\psi \vee \neg(\exists y)\neg\psi))$ of the Law of Excluded Middle and the sequent $((\forall y)\neg\neg\psi \vdash_{\vec{x}} (\forall y)\psi)$, which follows straightforwardly from the Law of Excluded Middle:

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{(\neg(\exists y)\neg\psi \vdash_{\vec{x}} \neg(\exists y)\neg\psi)}{((\neg(\exists y)\neg\psi \wedge (\exists y)\neg\psi) \vdash_{\vec{x}} \bot)}}{((\exists y)\neg\psi \vdash_{\vec{x}} \neg\neg(\exists y)\neg\psi)}}{(\neg\psi \vdash_{\vec{x},y} \neg\neg(\exists y)\neg\psi)}}{((\neg\psi \wedge \neg(\exists y)\neg\psi) \vdash_{\vec{x},y} \bot)}}{(\neg(\exists y)\neg\psi \vdash_{\vec{x},y} \neg\neg\psi)}}{(\neg(\exists y)\neg\psi \vdash_{\vec{x}} (\forall y)\neg\neg\psi) .}$$

Hence $(\phi \vdash \psi)$ is classically derivable in $\mathbb{T}$. □

Suggestions for further reading: Freyd [371], Freyd & Scedrov [381], Makkai & Reyes [790, 791].

D2

SKETCHES

D2.1 The concept of sketch

This chapter (rather like Chapter A3) is something of a digression from our main line of development: we shall not make any essential use of sketches in anything we wish to do with categorical logic, but, since they are extensively used by others in this context, it seems a good idea to give an account of what they are and how they relate to the line that we are following. Sketches were first introduced by C. Ehresmann [332] and were extensively developed by his pupils, notably L. Coppey and C. Lair [658]; for many years they were regarded as the 'exclusive property' of the French school of category-theorists (of which Ehresmann was the founder), but since the publication of the books by Barr and Wells [87, 88], and even more since the work of Makkai and Paré [787], their use has become much more widespread. In this account, we shall largely follow the conventions of Barr and Wells regarding terminology and notation.

Briefly, a sketch is an alternative way of 'presenting' a theory whose models we wish to study in appropriate categories, which is 'categorical' in character as compared with the 'linguistic' presentation of Chapter D1. A sketch is in some sense intermediate between a theory, as defined in 1.1.6, and its syntactic category as constructed in Section D1.4; it has some of the advantages of each, but also some of the disadvantages of each. Its principal advantage, as compared with a theory, is that we do not have to go through the tedious recursive construction of terms and formulae over a given signature before we can write down the axioms; the analogue of a 'signature' from which we start (the underlying directed graph of a sketch) contains enough to write down all the axioms directly in terms of it. Compared with a syntactic category, its main advantage is economy; for example, the syntactic category of even a simple theory such as that of groups is infinite, and quite complicated to describe, whereas the sketch for groups is finite and can be written down quite explicitly. On the other hand, it is appreciably *less* economical than the linguistic presentation of a theory; again taking groups as an example, the usual linguistic presentation requires one sort, three function symbols, no relation symbols and three axioms, whereas the sketch for groups has four objects, sixteen morphisms and nineteen axioms (as we shall see in 2.1.4(*b*) below). And sketches share with theories the disadvantage of 'signature-dependence': equivalence of sketches is a much stronger

condition than the assertion that they have equivalent categories of models in all appropriate categories.

To begin the formal definition, recall that a *directed graph* is a '(small) category without composition or identities'; that is, it has a set of *vertices* or objects, a set of *arrows* or morphisms, and a *source* and *target* (or domain and codomain) for each arrow. (To help us distinguish categories from directed graphs, we shall generally use the italicized terminology, which is not in common use for categories, when working in the latter.) Clearly, a directed graph is just a model of a certain two-sorted algebraic theory (cf. 1.1.7(b)); alternatively, it is a functor $\mathcal{C} \to \mathbf{Set}$, where $\mathcal{C}$ is the category with two objects and a parallel pair of morphisms between them (cf. A1.3.10(d)). The notion of *morphism* of directed graphs is the obvious one.

By a *diagram* in a directed graph G, we mean a graph morphism $d\colon G' \to G$; we say the diagram is *finite* if its domain is finite. If G is (the underlying directed graph of) a category, we say a diagram $d\colon G' \to G$ *commutes* if, for any two paths (= finite strings of arrows) in G' with the same source and target, the two morphisms in G obtained by composing the images under d of the arrows along the two paths are equal. (We allow the possibility that one of the paths might be empty; the composite of an empty path in a category is interpreted as being the appropriate identity morphism.) We say a diagram d is a *cone* if G' has a distinguished vertex g_0 such that, for any other vertex g, there is exactly one arrow $g_0 \to g$ in G' and no arrows $g \to g_0$. We say d is a *discrete cone* if it is a cone and there are no arrows in G' other than those with source g_0. If G' satisfies the corresponding conditions with 'source' and 'target' interchanged, we say d is a (*discrete*) *cocone*.

Definition 2.1.1

(a) A *sketch* $\mathbb{S}$ is a quadruple $(G_{\mathbb{S}}, D_{\mathbb{S}}, L_{\mathbb{S}}, C_{\mathbb{S}})$ where $G_{\mathbb{S}}$ is a directed graph, $D_{\mathbb{S}}$ is a set of finite diagrams in $G_{\mathbb{S}}$, $L_{\mathbb{S}}$ is a set of cones in $G_{\mathbb{S}}$ and $C_{\mathbb{S}}$ is a set of cocones in $G_{\mathbb{S}}$. (We normally omit the subscript $\mathbb{S}$'s, if there is no danger of confusion.)

(b) A *morphism of sketches* $f\colon \mathbb{S} \to \mathbb{T}$ is a morphism $f\colon G_{\mathbb{S}} \to G_{\mathbb{T}}$ of the underlying directed graphs, composition with which takes each member of $D_{\mathbb{S}}$ into a member of $D_{\mathbb{T}}$, each member of $L_{\mathbb{S}}$ into a member of $L_{\mathbb{T}}$ and each member of $C_{\mathbb{S}}$ into a member of $C_{\mathbb{T}}$. We write **Sk** for the category of sketches and sketch morphisms.

(c) A *model* of a sketch $\mathbb{S}$ in a category $\mathcal{C}$ is a graph morphism $M\colon G_{\mathbb{S}} \to \mathcal{C}$ such that

  (i) for each $d\colon G' \to G$ in $D_{\mathbb{S}}$, the composite $M \circ d$ commutes;

  (ii) for each $l\colon G' \to G$ in $L_{\mathbb{S}}$, the images under $M \circ l$ of the arrows with source g_0 in G' form a limit cone over the diagram formed by restricting $M \circ l$ to the subgraph G'' obtained by deleting g_0, and all arrows with source g_0, from G'; and

(iii) for each $c\colon G' \to G$ in $C_{\mathbb{S}}$, the images under $M \circ c$ of the arrows with target g_0 in G' similarly form a colimit cone under the rest of the diagram.

Equivalently (ignoring questions of size), a model of $\mathbb{S}$ in $\mathcal{C}$ is a sketch morphism from $\mathbb{S}$ to the 'underlying sketch of $\mathcal{C}$', where the latter is defined to be the underlying directed graph of $\mathcal{C}$, equipped with all finite diagrams which commute in $\mathcal{C}$, all limit cones in $\mathcal{C}$ and all colimit cones in $\mathcal{C}$. A *morphism of models* $f\colon M \to N$ is a natural transformation between graph morphisms (which makes sense because the codomain of M and N is not just a directed graph but a category). We write $\mathbb{S}$-$\mathbf{Mod}(\mathcal{C})$ for the category of $\mathbb{S}$-models in $\mathcal{C}$.

It will be noted that the definition of a model of a sketch does not presuppose any general information about the existence of limits or colimits in $\mathcal{C}$, merely that certain particular (co)cones are limits or colimits. However, when considering the category of $\mathbb{S}$-models in $\mathcal{C}$, it is customary to assume that $\mathcal{C}$ has limits and colimits of all the shapes which appear in $L_{\mathbb{S}}$ and $C_{\mathbb{S}}$.

Just as we classified theories according to the degree of linguistic complexity needed to formulate their axioms, so we classify sketches according to the types of cones and cocones which appear in them. We next list some of the classes that we may wish to consider.

Definition 2.1.2 Let $\mathbb{S} = (G, D, L, C)$ be a sketch.

(*a*) We say $\mathbb{S}$ is *linear* (or elementary) if both L and C are empty.

(*b*) We say $\mathbb{S}$ is a *limit sketch* (or projective sketch) if C is empty. If in addition all the cones in L are finite, we call $\mathbb{S}$ a *finite limit sketch* (or cartesian sketch). If C is empty and all the cones in L are discrete (and finite), we call $\mathbb{S}$ a (*finite*) *product sketch.*

(*c*) Similarly, we have the notions of (*finite*) *colimit sketch* (also called an inductive sketch) and of (*finite*) *coproduct sketch* when L is empty.

(*d*) The term *mixed sketch* is sometimes used (for emphasis) to indicate that both L and C are nonempty. We say $\mathbb{S}$ is *finitary* if (it is mixed and) all members of L and C are finite.

(*e*) We say $\mathbb{S}$ is a (*finitary*) *disjunctive sketch* if all cones in L are finite, and all cocones in C are (finite and) discrete.

(*f*) We say $\mathbb{S}$ is a *regular sketch* if L consists of finite cones, and all the cocones in C have the form

$$\begin{tikzcd} g \arrow[rr, shift left, "\alpha"] \arrow[rr, shift right, "\beta"'] \arrow[dr, "\gamma"'] & & h \arrow[dl, "\delta"] \\ & k & \end{tikzcd}$$

(with distinguished vertex k) where

$$\begin{tikzcd} g \arrow[r, "\alpha"] \arrow[d, "\beta"'] \arrow[dr, "\gamma"] & h \arrow[d, "\delta"] \\ h \arrow[r, "\delta"'] & k \end{tikzcd}$$

(with distinguished vertex g) is a cone in L.

(g) We say $\mathbb{S}$ is a *coherent sketch* if all cones in L are finite, and all cocones in C are either of the form described in (f), or finite and discrete.

(h) We say $\mathbb{S}$ is a *geometric sketch* if all cones in L are finite, but there are no restrictions on the cocones in C.

Remark 2.1.3 After seeing the definition of coherent sketch, the reader may have been surprised to discover that that of geometric sketch imposes no restrictions at all on cocones, rather than requiring the 'infinitary analogue' of the restrictions in 2.1.2(g). We could have done this, but the result would have been the same as far as the expressive power of the class of geometric sketches is concerned; essentially, the reason for this is contained in A1.4.19, where we saw that an ∞-pretopos has all small colimits. See also 2.2.8 below.

Of course, the reappearance in 2.1.2 of various words (regular, coherent, geometric, ...) that were used to describe classes of theories in Chapter D1 is no accident: we shall see in the next section that these classes of sketches have the same expressive power as the classes of theories with the same names. Before turning to this, however, we give some examples of the expressive power of the various classes of sketches listed above.

Examples 2.1.4 (a) If $\mathbb{S}$ is a linear sketch, then an $\mathbb{S}$-model in $\mathcal{C}$ is essentially the same thing as a functor from $\mathcal{C}_{\mathbb{S}} = (FG_{\mathbb{S}}/\equiv)$ to $\mathcal{C}$, where $FG_{\mathbb{S}}$ is the free category on $G_{\mathbb{S}}$ (i.e. the category whose objects are the vertices of $G_{\mathbb{S}}$ and whose morphisms are the paths in $G_{\mathbb{S}}$), and $\equiv$ is the smallest congruence on $FG_{\mathbb{S}}$ which identifies all parallel pairs which occur in the diagrams in $D_{\mathbb{S}}$. Conversely, given a small category $\mathcal{D}$, we can construct a linear sketch whose models are functors defined in $\mathcal{D}$, by taking G to be the underlying directed graph of $\mathcal{D}$, and D to consist of all commutative triangles

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

in $\mathcal{D}$ and all loops

$$A \circlearrowleft 1_A$$

where 1_A is an identity morphism in $\mathcal{D}$. Thus, up to Morita equivalence, linear sketches are essentially the same thing as small categories.

(*b*) As promised, we give a presentation of the sketch for groups, which is a finite product sketch. The underlying directed graph has four vertices which we denote by 1, G, G^2 and G^3 (anticipating what their interpretations will be in a model of the sketch) and the following sixteen arrows: $!\colon G \to 1$; $e\colon 1 \to G$; two arrows 1_G and $i\colon G \to G$; three arrows π_1, π_2 and $m\colon G^2 \to G$; two arrows $e \times 1$ and $(i,1)\colon G \to G^2$; three arrows π_1, π_2 and $\pi_3\colon G^3 \to G$; and four arrows $\pi_{12}, \pi_{23}, m \times 1$ and $1 \times m\colon G^3 \to G^2$. L consists of the three discrete cones

$$1 \qquad \begin{tikzcd} & G^2 \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ G & & G \end{tikzcd} \qquad \begin{tikzcd} & G^3 \arrow[dl, "\pi_1"'] \arrow[d, "\pi_2"] \arrow[dr, "\pi_3"] & \\ G & G & G \end{tikzcd}$$

(of which the first corresponds to the assertion that 1 is the limit of the empty diagram). D contains the loop diagram which says that 1_G is the identity on G, plus, for each of the six arrows with target G^2, a pair of diagrams specifying that (the interpretation in a model of) the arrow is 'what its name implies'. For example, the diagrams for $m \times 1$ are

$$\begin{tikzcd} G^2 \arrow[d, "m"'] & G^3 \arrow[l, "\pi_{12}"'] \arrow[d, "m \times 1"'] \arrow[dr, "\pi_3"] & \\ G & G^2 \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] & G \end{tikzcd}$$

and those for $(i,1)$ are

$$\begin{tikzcd} & G \arrow[dl, "i"'] \arrow[d, "{(i,1)}"] \arrow[dr, "1_G"] & \\ G & G^2 \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] & G \end{tikzcd} .$$

Finally, the three equations usually thought of as the axioms of group theory are represented by the diagrams

$$\begin{tikzcd} G^3 \arrow[r, "m \times 1"] \arrow[d, "1 \times m"] & G^2 \arrow[d, "m"'] & G \arrow[l, "e \times 1"'] \arrow[dl, "1_G"] \\ G^2 \arrow[r, "m"] & G & \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} G \arrow[r, "{(i,1)}"] \arrow[d, "!"'] & G^2 \arrow[d, "m"] \\ 1 \arrow[r, "e"] & G \end{tikzcd} \quad .$$

It is clear that any (many-sorted) algebraic theory, as defined in 1.1.7(a) and (b), can be presented in a similar way as a finite product sketch. Conversely, any finite product sketch $\mathbb{S}$ gives rise to a many-sorted algebraic theory: we take the vertices of $G_{\mathbb{S}}$ as sorts and its arrows as unary function symbols, plus, for each cone in $L_{\mathbb{S}}$ with arrows $(f_i\colon A \to B_i \mid 1 \leq i \leq n)$, an n-ary function symbol $g\colon B_1 \cdots B_n \to A$. Each 'commuting cell' in a diagram in $D_{\mathbb{S}}$ gives rise to an equation in the obvious way, and each n-ary cone gives rise to $n+1$ equations

$$(g(f_1(x), \ldots, f_n(x)) = x)$$

and $(f_i(g(y_1, \ldots, y_n)) = y_i)$ $(1 \leq i \leq n)$.

(c) Similarly, (infinite) product sketches can be used to describe 'infinitary algebraic theories' where one has operations acting on infinitely many elements of the structure. For example, for any infinite cardinal κ, one can formulate a sketch for κ-complete lattices, i.e. lattices in which every subset of cardinality less than κ has a join and a meet. (However, the smallness restrictions built into the definition of a sketch prevent us from describing a sketch for complete lattices, in which all subsets have joins and meets; this is demonstrated by the well-known fact that there is no free complete lattice on three generators [427]. See also 2.3.13 below.) These structures do not fit into the linguistic approach as we developed it in Chapter D1, since they require formulae with infinitely many free variables; but there is an evident extension of the notion of first-order language which would allow this.

(d) If $\mathcal{C}$ is a small cartesian category, there is a finite limit sketch whose models are cartesian functors defined on $\mathcal{C}$: just take the linear sketch for functors defined on $\mathcal{C}$, as given in (a), and add a limit cone for each finite diagram in $\mathcal{C}$ (equivalently, for the empty diagram and for each diagram formed by a pair of morphisms with common codomain). Thus, in view of 1.4.7, we know that every cartesian theory is Morita-equivalent to a finite limit sketch. (We shall prove the converse in 2.2.1 below.) Similarly, if $\mathcal{C}$ has (some) infinite limits, we can construct a limit sketch whose models are functors preserving these limits. In linguistic terms, such sketches correspond to the infinitary analogues of cartesian theories (which have been called 'ruled theories' by Isbell [475], modulo some differences about size restrictions); as in (c), these theories require an

infinitary extension of the language as we defined it in Section D1.1, but can be expressed by sequents built up using (finite and) infinite conjunction, and existential quantification subject to the restriction of 1.3.4.

(*e*) An example of a (finitary) disjunctive sketch is the sketch for fields: here we take a directed graph whose vertices, in addition to 1, F, F^2 and F^3, include an object U which is the distinguished vertex of a cone

$$\begin{tikzcd} & U \arrow[dl] \arrow[dr] \arrow[dd] & \\ F^2 \arrow[rr, "m"] \arrow[dr] & & F \\ & 1 \arrow[ur, "1"'] & \end{tikzcd}$$

(where m is the multiplication of F), whose effect is to ensure that, in any model M of the sketch, $M(U)$ is the 'object of units' of $M(F)$. Then the axioms which distinguish fields from commutative rings can be expressed by a discrete cocone

$$\begin{tikzcd} 1 \arrow[dr, "0"'] & & U \arrow[dl] \\ & F & \end{tikzcd}$$

(where the arrow $U \to F$ is not the same one that appears in the cone above, but is the composite of the arrow $U \to F^2$ with one of the projections).

(*f*) We shall give examples of regular and coherent sketches in the next section. A simple example of a finitary (mixed) sketch which is not coherent is given by the sketch for connected directed graphs. Directed graphs are of course models of the linear sketch corresponding to the directed graph

$$A \rightrightarrows V;$$

to express the idea of connectedness, we add a further vertex 1 with arrows $A \to 1$ and $V \to 1$, plus the discrete cone which says that 1 is a limit for the

empty diagram, and the cocone

$$\begin{tikzcd} A \arrow[rr, shift left] \arrow[rr, shift right] \arrow[dr] & & V \arrow[dl] \\ & 1 & \end{tikzcd} .$$

We shall see in 2.4.12(a) below that connected graphs cannot be the models of a finitary first-order theory; in particular, they are not the models of a coherent theory.

(g) Let $(\mathcal{C}, T)$ be a small site whose underlying category $\mathcal{C}$ is cartesian. Then we may construct a geometric sketch $\mathbb{S}$ whose models, in any geometric category $\mathcal{E}$, are cartesian cover-preserving functors $\mathcal{C} \to \mathcal{E}$ (where $\mathcal{E}$ is equipped with the coverage generated by small covering families); in particular, for a Grothendieck topos $\mathcal{E}$, we have

$$\mathbb{S}\text{-}\mathbf{Mod}(\mathcal{E}) \simeq \mathfrak{Top}\,(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, T))$$

by C2.3.9. Explicitly, we take the sketch for cartesian functors on $\mathcal{C}$, as defined in (d) above, and for each covering family $(f_i \colon A_i \to A \mid i \in I)$ we add a cocone whose vertices are A (the distinguished vertex), the A_i and the pullbacks $A_i \times_A A_j$ $((i,j) \in I \times I)$, and whose arrows are the f_i, the projections $A_i \times_A A_j \to A_i$ and $A_i \times_A A_j \to A_j$ and the composites $A_i \times_A A_j \to A_i \to A$. It is straightforward to verify that a cartesian functor $F \colon \mathcal{C} \to \mathcal{E}$ maps this cocone to a colimit iff the family $(F(f_i) \mid i \in I)$ is covering in $\mathcal{E}$.

(h) Let $(\mathcal{C}, T)$ be a site as in (g). We may also construct a limit sketch whose models (in **Set**) are sheaves on $\mathcal{C}$, as follows: we take the linear sketch for functors on $\mathcal{C}^{\mathrm{op}}$, as in ($a$) above, and then take each of the cocones described under (g), with all arrows reversed, as a distinguished cone of the sketch. Of course, this sketch will not be geometric unless all the covers in T can be generated by finite families; this will happen, for example, if $\mathcal{C}$ is a coherent category and T is its coherent coverage (cf. C2.1.12(d)).

Suggestions for further reading: Barr [77], Barr & Wells [87, 89], Makkai & Paré [787], Wells [1226].

D2.2 Sketches and theories

The next step in a conventional development of categorical logic in terms of sketches would be the construction, for each sketch $\mathbb{S}$ in one of the classes considered in 2.1.2, of a *classifying category* for $\mathbb{S}$ containing a model of $\mathbb{S}$ which is generic among $\mathbb{S}$-models in categories of the appropriate kind. However, since we have already done the corresponding job for theories in Section D1.4, we can by-pass this work by showing that sketches have the same expressive power as

theories, i.e. for every sketch in one of the appropriate classes there is a theory of the appropriate type having 'the same' models in all appropriate categories, and conversely. In fact we shall rely heavily on Section D1.4 in order to do this; we shall see how to convert sketches into Morita-equivalent theories directly, but in going the other way we shall appeal to the fact that (for example) any regular theory $\mathbb{T}$ is Morita-equivalent to the theory of regular functors on a small regular category $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$, and then show how the latter may be converted into a regular sketch.

For cartesian theories and finite limit sketches, the latter half of the work has been done already in 2.1.4(d):

Proposition 2.2.1 *Let* $\mathfrak{Cart}$ *denote the 2-category of small cartesian categories, cartesian functors and arbitrary natural transformations between them. For a 2-functor* $\mathbb{F}\colon \mathfrak{Cart} \to \mathfrak{Cat}$, *the following are equivalent:*

(i) $\mathbb{F}$ *is representable (in the 'up-to-equivalence' sense).*

(ii) *There exists a cartesian theory* $\mathbb{T}$ *such that* $\mathbb{F} \simeq \mathbb{T}\text{-}\mathbf{Mod}(-)$.

(iii) *There exists a finite limit sketch* $\mathbb{S}$ *such that* $\mathbb{F} \simeq \mathbb{S}\text{-}\mathbf{Mod}(-)$.

Proof (ii) $\Rightarrow$ (i) was proved in 1.4.7, and (i) $\Rightarrow$ (iii) in 2.1.4(d). So it suffices to prove (iii) $\Rightarrow$ (ii).

Given a finite limit sketch $\mathbb{S}$, we take a signature Σ whose sorts are the vertices of the graph $G_{\mathbb{S}}$, with the arrows of $G_{\mathbb{S}}$ as unary function symbols (and no relation symbols, other than equality). For each parallel pair of paths in a diagram in $D_{\mathbb{S}}$, we adopt an axiom of the form $(\top \vdash_x (s = t))$, where s and t are terms of the form $f(g(\cdots(x)\cdots))$. Now suppose we have a finite cone in $L_{\mathbb{S}}$ with vertices $A_0, A_1, \ldots, A_n$ (A_0 being the distinguished one), distinguished arrows $f_i\colon A_0 \to A_i$ ($1 \leq i \leq n$) and other arrows typically denoted by $g\colon A_i \to A_j$. Then we adopt the axioms $(\top \vdash_x (g(f_i(x)) = f_j(x)))$ for each such g, together with

$$\left(\bigwedge_{i=1}^{n}(f_i(x) = f_i(y)) \vdash_{x,y} (x = y)\right)$$

and

$$\left(\bigwedge(g(x_i) = x_j) \vdash_{x_1,\ldots,x_n} (\exists x_0) \bigwedge_{i=1}^{n}(f_i(x_0) = x_i)\right)$$

where the conjunction on the left-hand side of the last sequent is over all arrows g whose source is not A_0. Clearly, the second displayed sequent is cartesian relative to the first. Also, a Σ-structure in a (cartesian) category $\mathcal{C}$ is the same thing as a graph morphism $G_{\mathbb{S}} \to \mathcal{C}$; it satisfies the axioms of the first type iff it maps the diagrams in $D_{\mathbb{S}}$ to commutative diagrams in $\mathcal{C}$, and it satisfies those of the other three types iff it transforms the cones in $L_{\mathbb{S}}$ to limit cones in $\mathcal{C}$. □

Remark 2.2.2 Proposition 2.2.1 appears to state less than we have actually proved, in that it refers only to $\mathbb{T}$-models and $\mathbb{S}$-models in small cartesian categories, and we have considered them in arbitrary cartesian categories. But the difference is only apparent, since it is clear that, given any $\mathbb{T}$-model M in a cartesian category $\mathcal{C}$, we can find a small subcategory $\mathcal{C}'$ of $\mathcal{C}$, closed under finite limits, which contains M. Similar remarks apply to the results to be proved below about regular and coherent theories and sketches.

Proposition 2.2.3 *Let $\mathfrak{Reg}$ denote the 2-category of small regular categories, regular functors and arbitrary natural transformations between them. For a 2-functor $\mathbb{F}\colon \mathfrak{Reg} \to \mathfrak{Cat}$, the following are equivalent:*

(i) *$\mathbb{F}$ is representable.*

(ii) *There exists a regular theory $\mathbb{T}$ such that $\mathbb{F} \simeq \mathbb{T}$-$\mathbf{Mod}(-)$.*

(iii) *There exists a regular sketch $\mathbb{S}$ such that $\mathbb{F} \simeq \mathbb{S}$-$\mathbf{Mod}(-)$.*

Proof Once again, (ii) $\Rightarrow$ (i) was proved in Section D1.4.

(i) $\Rightarrow$ (iii): Let $\mathcal{C}$ be a small regular category. We construct a regular sketch $\mathbb{S}$ whose models are regular functors defined on $\mathcal{C}$, by taking the finite limit sketch of 2.1.4(d) and adding the cocone

$$\begin{tikzcd} R \arrow[rr, shift left, "a"] \arrow[rr, shift right, "b"'] \arrow[dr, "fa"'] & & A \arrow[dl, "f"] \\ & B & \end{tikzcd}$$

whenever f is a cover in $\mathcal{C}$ and (a, b) is its kernel-pair. (This is allowable under 2.1.2(f), since the cones of $\mathbb{S}$ include that for the pullback of f against itself.) Clearly, a model of $\mathbb{S}$ is a cartesian functor defined on $\mathcal{C}$ which sends covers to coequalizers of their kernel-pairs; but this is the same thing as a regular functor, by A1.3.4.

(iii) $\Rightarrow$ (ii): Given a regular sketch $\mathbb{S}$, we construct a regular theory $\mathbb{T}$ with the same models, by taking the cartesian theory described in the proof of 2.2.1 and adding the sequent $(\top \vdash_y (\exists x\colon A)(f(x) = y))$ for each cocone

$$\begin{tikzcd} R \arrow[rr, shift left, "a"] \arrow[rr, shift right, "b"'] \arrow[dr, "g"'] & & A \arrow[dl, "f"] \\ & B & \end{tikzcd}$$

in $\mathbb{S}$. Then a model of $\mathbb{T}$ in a regular category $\mathcal{D}$ is a model M of the finite limit part of $\mathbb{S}$, such that for each cocone as above the morphism $M(f)$ is a cover

in $\mathcal{D}$. But since the finite limit part of $\mathbb{S}$ includes the information that (a, b) is the kernel-pair of f, this is equivalent to saying that $M(f)$ is the coequalizer of $M(a)$ and $M(b)$, i.e. that M is a model of $\mathbb{S}$. □

Remark 2.2.4 Proposition 2.2.3 remains valid if we replace the 2-category $\mathfrak{Reg}$ by the sub-2-category $\mathfrak{EffReg}$ of effective regular categories, by 1.4.12(i).

To get a similar result for the coherent case, we actually have to restrict our attention to models in positive coherent categories; this is a consequence of the way in which we defined coherent sketches, which allows us to talk about disjoint unions but not about arbitrary (finite) unions. It would be possible to formulate a more 'relaxed' version of Definition 2.1.2(g) which would not have this disadvantage (cf. the proof of 2.2.7(ii) below); but it would be appreciably more complicated. And, since we are primarily interested in models in toposes, the restriction is not one that worries us unduly; in fact we shall formulate the next proposition only for models in pretoposes.

Proposition 2.2.5 *Let $\mathfrak{PreTop}$ denote the 2-category of small pretoposes, coherent functors and arbitrary natural transformations between them. For a 2-functor $\mathbb{F}\colon \mathfrak{PreTop} \to \mathfrak{Cat}$, the following are equivalent:*

(i) *$\mathbb{F}$ is representable.*

(ii) *There exists a coherent theory $\mathbb{T}$ such that $\mathbb{F} \simeq \mathbb{T}$-$\mathbf{Mod}(-)$.*

(iii) *There exists a coherent sketch $\mathbb{S}$ such that $\mathbb{F} \simeq \mathbb{S}$-$\mathbf{Mod}(-)$.*

Proof (ii) ⇒ (i) was proved in 1.4.12(ii). For (i) ⇒ (iii), let $\mathcal{C}$ be a small pretopos; take the regular sketch constructed in 2.2.3, and add the discrete cocone which says that 0 is an initial object, and the cocones

$$\begin{tikzcd} A \arrow[dr, "\nu_1"'] & & B \arrow[dl, "\nu_2"] \\ & A+B & \end{tikzcd}$$

for each pair of objects (A, B). A model of this sketch is then a regular functor defined on $\mathcal{C}$ which preserves finite coproducts; but since unions can be constructed from coproducts and images, this is the same thing as a coherent functor.

(iii) ⇒ (ii): Given a coherent sketch $\mathbb{S}$, take the theory constructed in the proof of 2.2.3 from the regular part of $\mathbb{S}$, and add the sequents

$$\left(\top \vdash_x \bigvee_{i=1}^{n} (\exists x_i)(f_i(x_i) = x)\right),$$

$((f_i(x_i) = f_i(x'_i)) \vdash_{x_i, x'_i} (x_i = x'_i))$ (for each i) and $((f_i(x_i) = f_j(x_j)) \vdash_{x_i, x_j} \bot)$ (for each $i \neq j$) for every discrete cocone with arrows $f_i\colon A_i \to A$ $(1 \leq i \leq n)$.

Then a $\mathbb{T}$-model in a pretopos $\mathcal{D}$ is a model of the regular part of $\mathbb{S}$, which sends each discrete cocone in $\mathbb{S}$ to a disjoint union of subobjects (equivalently, a coproduct) in $\mathcal{D}$; so it is the same thing as an $\mathbb{S}$-model in $\mathcal{D}$. □

Remark 2.2.6 It will be observed that the sequents constructed from discrete cocones in the proof of 2.2.5(iii) ⇒ (ii) form a disjunctive theory, as defined in 1.3.6; this forms part of the proof (which we shall not give in detail) that disjunctive theories have the same expressive power as disjunctive sketches in the sense of 2.1.2(*e*).

For geometric theories, we cannot state the analogue of 2.2.1, 2.2.3 and 2.2.5 in precisely the same form, because of the fact that ∞-pretoposes are generally not small. Nevertheless, we still have

Proposition 2.2.7

(i) *Let $\mathbb{S}$ be a geometric sketch. Then there exists a geometric theory $\mathbb{T}$ such that, for all ∞-pretoposes $\mathcal{C}$, we have $\mathbb{S}$-$\mathbf{Mod}(\mathcal{C}) \simeq \mathbb{T}$-$\mathbf{Mod}(\mathcal{C})$, naturally in $\mathcal{C}$.*

(ii) *Let $\mathbb{T}$ be a geometric theory. Then there exists a geometric sketch $\mathbb{S}$ such that, for all geometric categories $\mathcal{C}$, we have $\mathbb{T}$-$\mathbf{Mod}(\mathcal{C}) \simeq \mathbb{S}$-$\mathbf{Mod}(\mathcal{C})$, naturally in $\mathcal{C}$.*

Proof (i) Given $\mathbb{S}$, we construct the cartesian theory corresponding to the finite limit part of $\mathbb{S}$, as in 2.2.1. Then, for each cocone in $\mathcal{C}_{\mathbb{S}}$ with distinguished arrows $(f_i \colon A_i \to A \mid i \in I)$, we add the sequent

$$(\top \vdash_x \bigvee_{i \in I} (\exists x_i)(f_i(x_i) = x))$$

plus, for each non-distinguished arrow $g \colon A_i \to A_j$, the sequent

$$(\top \vdash_{x_i} (f_j(g(x_i)) = f_i(x_i)))$$

and, for each pair (i, j), the sequent

$$\left((f_i(x_i) = f_j(x_j)) \vdash_{x_i, x_j} \bigvee (\exists y_1, \ldots, y_{n+1})\left(\bigwedge_{k=1}^{n} \phi_k \wedge (x_i = y_1) \wedge (x_j = y_{n+1})\right)\right)$$

where the disjunction on the right is over all finite chains of non-distinguished arrows (in either direction)

$$A_i \xrightarrow{\quad g_1 \quad} \bullet \xrightarrow{\quad g_2 \quad} \bullet \quad \cdots \quad \bullet \xrightarrow{\quad g_n \quad} A_j$$

and ϕ_k is either $(g_k(y_k) = y_{k+1})$ or $(g_k(y_{k+1}) = y_k)$, depending on the direction of g_k. It is straightforward to verify that, if M is a model of the finite limit part of $\mathbb{S}$ in an ∞-pretopos $\mathcal{C}$, then M satisfies the sequents displayed above iff MA

is covered by the coproduct of the MA_i, $i \in I$, and the kernel-pair of this cover is exactly the equivalence relation by which we have to factor the coproduct to obtain the colimit of the diagram formed by (the images under M of) the non-distinguished arrows in the cocone. (Compare the construction of arbitrary coequalizers in an ∞-pretopos, given in A1.4.19.) Thus a model of the theory described above is a model of $\mathbb{S}$.

(ii) Given $\mathbb{T}$, the geometric category $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ is generally not small, but it is equivalent to a small category, since every object is isomorphic to one of the form $\{\vec{x}.\phi\}$ where ϕ is a disjunction of regular formulae, by 1.3.8(ii). So it suffices to prove that, given a small geometric category $\mathcal{C}$, we can construct a geometric sketch $\mathbb{S}$ whose models correspond to geometric functors defined on $\mathcal{C}$. For this, we take the regular sketch constructed in 2.2.3, and for each union of subobjects $A = \bigcup_{i \in I} A_i$ in $\mathcal{C}$ we add a cocone with distinguished vertex A, whose other vertices are all the A_i and the pairwise intersections $A_i \cap A_j$, and whose arrows are all the inclusions $A_i \rightarrowtail A$, $A_i \cap A_j \rightarrowtail A$, $A_i \cap A_j \rightarrowtail A_i$ and $A_i \cap A_j \rightarrowtail A_j$. By the infinitary analogue of A1.4.3, a functor $F: \mathcal{C} \to \mathcal{D}$ (where $\mathcal{D}$ is geometric) maps this cocone to a colimit iff FA is the union of the FA_i. Hence F is a model of the sketch just described iff it is a regular functor and preserves arbitrary unions, i.e. iff it is a geometric functor. $\square$

Remark 2.2.8 As we remarked in 2.1.3, our definition of geometric sketch in 2.1.2(h) is not simply the infinitary analogue of 2.1.2(g), in that it includes no restrictions on the types of cocones in $\mathbb{S}$. The proof of 2.2.7(i) provides the justification for this choice; in a sense, it is just a reworking of the observation, already made in A1.4.19, that an ∞-pretopos has all small colimits. Note also that it follows trivially from 2.2.7 that geometric sketches and geometric theories have the same expressive power in ∞-pretoposes; however, the use of arbitrary geometric categories, rather than ∞-pretoposes, in the proof of (ii) was essential, since an ∞-pretopos (has all small coproducts, and hence) is almost never small.

The reader may also be wondering whether it is possible to characterize a class of theories having the same expressive power as the class of (arbitrary) finitary sketches. At least as far as models in Grothendieck toposes are concerned, the answer is yes, thanks to a result which we noted in Section A2.5. In order to present it, let us define a *σ-coherent theory* to be a geometric theory whose axioms involve only (finite or) countable disjunctions. We shall also call a sketch σ-coherent if it contains only finite cones and (finite or) countable cocones.

Proposition 2.2.9

(i) *Let $\mathbb{S}$ be a σ-coherent sketch. Then there exists a σ-coherent theory $\mathbb{T}$ such that, for all σ-pretoposes $\mathcal{C}$, we have $\mathbb{S}$-$\mathbf{Mod}(\mathcal{C}) \simeq \mathbb{T}$-$\mathbf{Mod}(\mathcal{C})$, naturally in $\mathcal{C}$.*

(ii) *Let $\mathbb{T}$ be a σ-coherent theory. Then there exists a finitary sketch $\mathbb{S}$ such that, for all countably cocomplete toposes $\mathcal{E}$, we have $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \simeq \mathbb{S}$-$\mathbf{Mod}(\mathcal{E})$, naturally in $\mathcal{E}$.*

Proof (i) follows immediately from an inspection of the proof of 2.2.7(i): if the cocones in $\mathbb{S}$ are all countable, then so are all the disjunctions in the sequents displayed there.

(ii) Given a σ-coherent theory $\mathbb{T}$, we may construct a (small) σ-coherent syntactic category $\mathcal{C}_{\mathbb{T}}^{\sigma\mathrm{coh}}$ in the usual way, and then positivize and effectivize it to a σ-pretopos $\mathcal{P}_{\mathbb{T}}^{\sigma}$, such that for any σ-pretopos $\mathcal{E}$ (in particular, for any topos with countable coproducts) we have

$$\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E}) \simeq \sigma\text{-}\mathfrak{PreTop}(\mathcal{P}_{\mathbb{T}}^{\sigma}, \mathcal{E})$$

(this is just the 'sigma' analogue of 1.4.12(ii) and (iii)). But if $\mathcal{E}$ is a topos, then by A2.5.7 the objects of the category on the right-hand side (the σ-coherent functors $\mathcal{P}_{\mathbb{T}}^{\sigma} \to \mathcal{E}$) are exactly those functors which are both cartesian and cocartesian. So we may identify them with $\mathbb{S}$-models in $\mathcal{E}$, where $\mathbb{S}$ is the sketch formed by the underlying directed graph of $\mathcal{P}_{\mathbb{T}}^{\sigma}$, together with all commutative diagrams, all finite limit cones and all finite colimit cocones in this category. □

Thus the expressive power of finitary sketches coincides with that of σ-coherent theories, and with that of σ-coherent sketches. In particular, we note that the theory of K-finite objects, introduced in 1.1.7(k), is sketchable by a finitary sketch, although (as we shall see in 2.4.12(c) below) it is not Morita-equivalent to any finitary theory.

Suggestions for further reading; Adámek et al. [19], Makkai [785].

D2.3 Sketchable and accessible categories

One of the most remarkable results in the theory of sketches is that the class of sketchable categories – that is, those categories which occur as $\mathbb{S}$-$\mathbf{Mod}(\mathbf{Set})$ for some sketch $\mathbb{S}$ – can be characterized up to equivalence in purely categorical terms, as the class of accessible categories. We shall not give the proof of this result in full detail, since it is somewhat peripheral to our main concerns, and it is adequately covered in the books by Makkai and Paré [787] and by Adámek and Rosický [23]. Nevertheless, it seems appropriate to give at least a rough outline of the argument.

We begin with some terminology. Throughout this section we shall be working with locally small categories (that is, categories enriched over a fixed category $\mathbf{Set}$ of sets), and the letters κ and λ will denote (infinite) regular cardinals in $\mathbf{Set}$. A category $\mathcal{C}$ is called *κ-small* if the class of all its morphisms is a set of cardinality less than κ, and $\mathcal{C}$ is called *κ-filtered* if every κ-small diagram in $\mathcal{C}$ (that is, every functor $\mathcal{D} \to \mathcal{C}$ where $\mathcal{D}$ is κ-small) has a cone under it. (This is in particular the case when $\mathcal{C}$ has all κ-small colimits, in which case we call it *κ-cocomplete.*)

The usual notion of filteredness, studied in Section B2.6 for internal categories, is equivalent to ω-filteredness, where ω is the cardinality of the set of

natural numbers. For a general κ, κ-filteredness can also be expressed in more 'elementary' terms, similar to the conditions of B2.6.2: it is sufficient to consider discrete diagrams (i.e. families of objects) of cardinality less than κ, and diagrams consisting of two objects and a family of fewer than κ parallel morphisms between them. We shall also need the 'κ-analogue' of (the specialization to **Set** of) the main result of Section B2.6: namely, that a small category $\mathcal{C}$ is κ-filtered iff the colimit functor $\varinjlim_{\mathcal{C}} : [\mathcal{C}, \mathbf{Set}] \to \mathbf{Set}$ preserves all κ-small limits. The proof is a straightforward generalization of the familiar one for $\kappa = \omega$.

Definition 2.3.1 Let $\mathcal{C}$ be a locally small category with κ-filtered colimits (that is, colimits over all small κ-filtered categories).

(*a*) An object A of $\mathcal{C}$ is called *κ-presentable* if the functor $\mathcal{C}(A, -) : \mathcal{C} \to \mathbf{Set}$ preserves κ-filtered colimits. (In the special case $\kappa = \omega$, we normally say 'finitely presentable' rather than 'ω-presentable'.)

(*b*) $\mathcal{C}$ is called *κ-accessible* if it contains a set $\mathcal{G}$ of κ-presentable objects, such that every object of $\mathcal{C}$ is expressible as a κ-filtered colimit of objects in $\mathcal{G}$.

A category is called *accessible* if there exists a regular cardinal κ for which it (has κ-filtered colimits and) is κ-accessible. It is clear that a κ-presentable object is λ-presentable for all regular $\lambda > \kappa$; it is not true in general that a κ-accessible category is λ-accessible for all $\lambda > \kappa$, but for any accessible category $\mathcal{C}$ there exists a cardinal λ such that $\mathcal{C}$ is λ'-accessible for all regular $\lambda' > \lambda$.

The set $\mathcal{G}$ appearing in part (*b*) of the definition is clearly a generating set for $\mathcal{C}$, as defined in Section A1.2; but the condition imposed in (*b*) is stronger than this. There is (up to equivalence) a 'canonical' choice for $\mathcal{G}$:

Lemma 2.3.2 *Let $\mathcal{C}$ be a κ-accessible category. Then*

(i) *The full subcategory $\mathcal{C}_\kappa$ of κ-presentable objects of $\mathcal{C}$ is essentially small (i.e. equivalent to a small category).*

(ii) *For any object A of $\mathcal{C}$, the comma category $(\mathcal{C}_\kappa \downarrow A)$ (whose objects are morphisms $B \to A$ with $B \in \mathrm{ob}\, \mathcal{C}_\kappa$, and whose morphisms are commutative triangles) is κ-filtered.*

(iii) *$\mathcal{C}_\kappa$ is* dense *in $\mathcal{C}$; i.e., for any object A of $\mathcal{C}$, the canonical cone under the forgetful functor $(\mathcal{C}_\kappa \downarrow A) \to \mathcal{C}$ with vertex A is colimiting.*

Proof If A is a κ-presentable object of $\mathcal{C}$, then by expressing it as a κ-filtered colimit of objects G_i in the set $\mathcal{G}$ given by 2.3.1(*b*), we see that the identity morphism $A \to A$ must factor through some $G_i \to A$, i.e. A is a retract of some G_i. Since each G_i has only a set of idempotent endomorphisms, it follows that there is only a set of isomorphism classes of κ-presentable objects of $\mathcal{C}$. So (i) is established. (Conversely, any retract of a κ-presentable object is easily seen to be κ-presentable; so $\mathcal{C}_\kappa$ is precisely the full subcategory of $\mathcal{C}$ on those objects which are retracts of members of $\mathcal{G}$.)

Now let A be an arbitrary object of $\mathcal{C}$, and let $D: J \to \mathcal{C}$ be a diagram in $\mathcal{C}$ with colimit A, where J is κ-filtered and the objects $D(j)$, $j \in \mathrm{ob}\, J$, are in $\mathcal{G}$. We have an obvious comparison functor $F: J \to (\mathcal{C}_\kappa \downarrow A)$ sending j to the jth leg of the colimiting cone; and from the very definition of κ-presentability, it is easy to see that this functor is final (that is, each object of $(\mathcal{C}_\kappa \downarrow A)$ admits an 'essentially unique' map to an object in the image of F). From this, it follows by standard arguments that, for any functor $G: (\mathcal{C}_\kappa \downarrow A) \to \mathcal{E}$ where $\mathcal{E}$ is cocomplete, the canonical morphism $\varinjlim_J (G \circ F) \to \varinjlim_{(\mathcal{C}_\kappa \downarrow A)}(G)$ is an isomorphism. Taking $\mathcal{E} = \mathbf{Set}$ and using the characterization of κ-filtered categories mentioned earlier, we deduce that $(\mathcal{C}_\kappa \downarrow A)$ is κ-filtered; and taking G to be the forgetful functor $(\mathcal{C}_\kappa \downarrow A) \to \mathcal{C}$ we obtain (iii), since the composite of this functor with F is simply D. □

Thus, given that $\mathcal{C}$ is κ-accessible, we may take the witnessing set $\mathcal{G}$ to be any representative set of objects of the subcategory $\mathcal{C}_\kappa$, and we shall normally do this from now on.

In general, the denseness of a small full subcategory $\mathcal{D}$ of $\mathcal{C}$ is equivalent to saying that the canonical functor $V: \mathcal{C} \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ which sends A to (the restriction to $\mathcal{D}$ of) the functor $\mathcal{C}(-, A)$ is full and faithful: for a morphism $V(A) \to V(B)$ in $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ is essentially the same thing as a cone under the forgetful functor $(\mathcal{D} \downarrow A) \to \mathcal{C}$ with vertex B. (Recall that ob $\mathcal{D}$ is a separating (resp. generating) set for $\mathcal{C}$ iff V is faithful (resp. conservative).)

This leads to an alternative characterization of accessible categories as subcategories of functor categories. We shall call a functor $F: \mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ a *κ-torsor* on $\mathcal{D}$ if it is expressible as a κ-filtered colimit of representable functors; this is equivalent to saying that the domain $\mathcal{F}$ of the discrete fibration $\mathcal{F} \to \mathcal{D}$ corresponding to F is κ-filtered (cf. B3.2.3). Using this, one may show that the category κ-$\mathbf{Tors}(\mathcal{D})$ of κ-torsors on $\mathcal{D}$ is closed under κ-filtered colimits in $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$.

Proposition 2.3.3 *A category $\mathcal{C}$ is κ-accessible iff it is equivalent to the category κ-$\mathbf{Tors}(\mathcal{D})$ for some small category $\mathcal{D}$.*

Proof One direction follows from 2.3.2: if $\mathcal{C}$ is κ-accessible, we take $\mathcal{D}$ to be a (small) skeleton of $\mathcal{C}_\kappa$, and then the functor $V: \mathcal{C} \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ induces one half of the required equivalence. Conversely, we must show that κ-$\mathbf{Tors}(\mathcal{D})$ is always κ-accessible. We have observed above that it has κ-filtered colimits; and since these are preserved by the inclusion κ-$\mathbf{Tors}(\mathcal{D}) \to [\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$, the Yoneda lemma ensures that the representable functors $\mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ are κ-presentable as objects of κ-$\mathbf{Tors}(\mathcal{D})$. So they will serve as the members of the set $\mathcal{G}$ required for 2.3.1(b). □

We remark that the κ-presentable objects of κ-$\mathbf{Tors}(\mathcal{D})$ are exactly the retracts of representable functors, by the argument in the proof of the first part of 2.3.2. (The 2-element monoid $\{1, e\}$ with $e^2 = e$ is a κ-filtered category for

any κ, so retracts of representable functors are always κ-torsors.) Thus, provided $\mathcal{D}$ is Cauchy-complete, we can recover it (up to equivalence) from κ-**Tors**$(\mathcal{D})$.

An important special case of 2.3.3 occurs when the category under consideration is (co)complete:

Proposition 2.3.4 *For a category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ is κ-accessible and cocomplete.*

(ii) *$\mathcal{C}$ is κ-accessible and κ-cocomplete.*

(iii) *$\mathcal{C}$ is equivalent to the category κ-**Cts** $(\mathcal{D}^{\mathrm{op}}, \mathbf{Set})$ of κ-continuous functors $\mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ (that is, functors preserving κ-small limits), where $\mathcal{D}$ is a κ-cocomplete small category.*

(iv) *$\mathcal{C}$ is κ-accessible and complete.*

Proof (i)$\Rightarrow$(ii) is trivial. If (ii) holds, it follows easily from the fact that κ-small limits commute with κ-filtered colimits in **Set** that $\mathcal{C}_\kappa$ is closed under κ-small colimits in $\mathcal{C}$; in particular it is κ-cocomplete (and so is any skeleton $\mathcal{D}$). But it is easy to see that any κ-torsor on $\mathcal{D}$ preserves κ-small limits, since it is a κ-filtered colimit of functors preserving all limits; and conversely if $\mathcal{D}$ is κ-cocomplete, then any κ-continuous functor is a κ-torsor, by the argument of B3.2.5. So condition (iii) follows from 2.3.3. For (iii)$\Rightarrow$(iv), note that κ-**Cts** $(\mathcal{D}^{\mathrm{op}}, \mathbf{Set})$ is closed under arbitrary small limits in $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$, since limits commute with limits.

It remains to show that (iv) implies (i). For this, we shall employ the General Adjoint Functor Theorem, to construct a left adjoint for the diagonal functor $\Delta \colon \mathcal{C} \to [J, \mathcal{C}]$ for any small category J. It is clear that (iv) implies that $\mathcal{C}$ has, and Δ preserves, all small limits, so we have to verify the solution set condition for a typical object $D \colon J \to \mathcal{C}$ of $[J, \mathcal{C}]$. Note first that since every object A of $\mathcal{C}$ is expressible as a small (κ-filtered) colimit of objects which are κ-presentable (and hence λ-presentable for any $\lambda > \kappa$), we can find λ such that A is λ-presentable. Hence, as J is small, we can find λ such that J is λ-small, $\mathcal{C}$ is λ-accessible, and all the objects in the image of D are λ-presentable. Now let $\mathcal{G}_\lambda$ be a representative set of λ-presentable objects of $\mathcal{C}$; then it is easy to see that the set of all cones under D with vertices in $\mathcal{G}_\lambda$ is a solution set for Δ at the object D. □

A direct proof of (iii)$\Rightarrow$(i) in 2.3.4 may be given by using the Adjoint Functor Theorem to show that κ-**Cts** $(\mathcal{D}^{\mathrm{op}}, \mathbf{Set})$ is reflective in $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$, for any κ; see [587].

Categories satisfying the equivalent conditions of 2.3.4 are called *locally κ-presentable* (or simply *locally presentable*, if we do not wish to specify the cardinal κ). They have been studied for considerably longer than general accessible categories; see [386]. In contrast to the situation noted after 2.3.1, a locally κ-presentable category $\mathcal{C}$ is always locally λ-presentable for all regular $\lambda > \kappa$; for the full subcategory $\mathcal{C}_\lambda$ is closed under λ-small colimits, and hence $(\mathcal{C}_\lambda \downarrow A)$ is λ-filtered (indeed, λ-cocomplete) for all A. However, in other respects the theory

of locally presentable categories generalizes with remarkably little difficulty to accessible categories.

Remark 2.3.5 Although, as we indicated at the start of this section, our concern is really with *infinite* regular cardinals κ, it is instructive to consider what happens to the results above if we take κ to be the (trivially regular) cardinal 0. Since no categories are 0-small, we see that every category is 0-filtered, and so an object A of a cocomplete category $\mathcal{C}$ is 0-presentable iff $\mathcal{C}(A,-)$ preserves arbitrary small colimits. Since a 0-accessible category is necessarily cocomplete, we thus deduce that *a category $\mathcal{C}$ is 0-accessible iff it is locally 0-presentable, iff it is equivalent to* $[\mathcal{D}^{\mathrm{op}}, \mathbf{Set}]$ *for some small category $\mathcal{D}$*. Note also that, if $\mathcal{C}$ is an ∞-pretopos, then an object A of $\mathcal{C}$ is 0-presentable iff $\mathcal{C}(A,-)$ is a geometric functor, iff A is an indecomposable projective (cf. A1.1.10); thus the result just quoted yields another proof of the characterization of functor categories amongst Grothendieck toposes given in C2.2.20.

Locally presentable categories correspond, in a fairly precise sense, to limit sketches:

Theorem 2.3.6 *For a category $\mathcal{C}$, the following are equivalent:*

(i) *$\mathcal{C}$ is locally κ-presentable.*

(ii) *There exists a limit sketch $\mathbb{S}$, in which all the distinguished cones have cardinality less than κ, such that $\mathcal{C} \simeq \mathbb{S}$-$\mathbf{Mod}(\mathbf{Set})$.*

Proof Suppose (i) holds. We construct a limit sketch $\mathbb{S} = (G, D, L)$ as follows: G is the underlying directed graph of (a small skeleton of) $(\mathcal{C}_\kappa)^{\mathrm{op}}$, D consists of all finite commutative diagrams in this category, and L consists of all limit cones of cardinality less than κ. Then models of $\mathbb{S}$ are precisely κ-continuous functors $(\mathcal{C}_\kappa)^{\mathrm{op}} \to \mathbf{Set}$, so by 2.3.4(iii) they form a category equivalent to $\mathcal{C}$.

For the converse, consider first the case $\kappa = \omega$. Then $\mathbb{S}$ is a finite limit sketch, so it corresponds to a cartesian theory $\mathbb{T}$ by 2.2.1, and its models can be identified with cartesian (that is, ω-continuous) functors $\mathcal{C}_\mathbb{T} \to \mathbf{Set}$. So by 2.3.4(iii) $\mathbb{S}$-$\mathbf{Mod}(\mathbf{Set})$ is locally finitely presentable. For larger values of κ, the argument is similar: we identify $\mathbb{S}$ with an appropriate theory in 'κ-cartesian logic', that is the fragment of infinitary logic in which we allow conjunctions of cardinality less than κ, and existential quantification subject to the restriction in 1.3.4, and then we generalize the construction of the syntactic category $\mathcal{C}_\mathbb{T}$ in 1.4.1, by allowing 'contexts' to be families of variables of cardinality less than κ. In this way we obtain a κ-complete small category $\mathcal{D}$ such that models of $\mathbb{S}$ correspond to κ-continuous functors defined on $\mathcal{D}$. □

Since we are concerned mainly with sketches in which all cones are finite, we shall generally be interested only in the case $\kappa = \omega$ of the above theorem.

However, it is worth pausing to note one consequence of the general case:

Corollary 2.3.7 *Any Grothendieck topos is locally presentable.*

Proof We saw in 2.1.4(h) that, for any small site $(\mathcal{C}, T)$ whose underlying category $\mathcal{C}$ is cartesian, there is a limit sketch whose models are T-sheaves on $\mathcal{C}$. But any Grothendieck topos can be presented as the category of sheaves on such a site, by C2.2.8(iii), so the result is immediate from 2.3.6. □

Generalizing the first half of the proof of 2.3.6, we have

Proposition 2.3.8 *Any accessible category is sketchable. More precisely, if $\mathcal{C}$ is κ-accessible, then there is a sketch $\mathbb{S}$ such that $\mathcal{C} \simeq \mathbb{S}$-**Mod**(**Set**), and all the cones in $\mathbb{S}$ have cardinality less than κ (although the cocones may be larger).*

Proof Let $\mathcal{D}$ be a small category such that $\mathcal{C} \simeq \kappa$-**Tors**$(\mathcal{D})$, as in 2.3.3. We shall construct a sketch whose models are (equivalent to) κ-torsors on $\mathcal{D}$.

Let $F \colon \mathcal{D}^{\mathrm{op}} \to \mathbf{Set}$ be a functor, and $\mathcal{F} \to \mathcal{D}$ the corresponding discrete fibration. The assertion that $\mathcal{F}$ satisfies the first of the 'elementary' conditions for κ-filteredness, mentioned before 2.3.1, is tantamount to saying that, for each family $(A_i \mid i \in I)$ of fewer than κ objects of $\mathcal{D}$, the canonical map

$$\coprod F(B) \longrightarrow \prod_{i \in I} F(A_i),$$

where the coproduct is over all families of morphisms $(f_i \colon A_i \to B \mid i \in I)$ with common codomain, is epic. But we may construct a sketch whose models are functors satisfying this condition, starting from the linear sketch for functors defined on $\mathcal{D}^{\mathrm{op}}$, by adding new vertices and arrows to the graph for the product and coproduct above, their projections and coprojections, the canonical morphism between them, and the identity morphism on $\prod_{i \in I} F(A_i)$, together with the diagrams which say that the two latter morphisms are what they ought to be, the cone and cocone which say that the product and coproduct are what they ought to be, and the cocone which says that the pushout of the displayed morphism above against itself is the identity. If we do this for all families of objects $(A_i \mid i \in I)$ of cardinality less than κ in $\mathcal{D}$, and perform a similar construction for the second 'elementary' condition for filteredness, we obtain the required sketch. □

In particular, we note that any finitely accessible category is sketchable by a geometric sketch; equivalently, by 2.2.7, it is (equivalent to) the category of models in **Set** of a geometric theory. However, the converse of 2.3.8 is not true in this 'sharp' form: a bound for the cardinalities of the cones in a sketch does not necessarily yield an 'index of accessibility' for its category of models.

Example 2.3.9 Consider the theory $\mathbb{T}$ of directed graphs in which every vertex is the source of some arrow. This is clearly a regular theory, and so is sketchable

by a finitary (indeed, finite) sketch. However, we shall show that its category of models is not ω-accessible. Let $N_<$ be the graph whose vertices are the natural numbers, with one arrow $m \to n$ whenever $m < n$, and $N_\leq$ the graph obtained from it by adding an arrow $n \to n$ for each n; clearly, both $N_<$ and $N_\leq$ are models of $\mathbb{T}$. If A is a nonempty finitely presentable $\mathbb{T}$-model and there is a homomorphism $f: A \to N_<$, then its composite with the inclusion $N_< \rightarrowtail N_\leq$ is also a homomorphism; but $N_\leq$ is the directed union of its finite full subgraphs (all of which are models of $\mathbb{T}$), and so the finite presentability of A ensures that f has finite image. But this means that there is a vertex v of A such that $f(v)$ is an upper bound for the image of f, which contradicts the requirement that there must exist some arrow $v \to w$ in A. So there can be no such f; hence $N_<$ cannot be expressed as a filtered colimit of finitely presentable $\mathbb{T}$-models.

In proving the converse of 2.3.8, the key ingredient is a 'downward Löwenheim–Skolem theorem' for models of sketches. Given a **Set**-model M of a (small) sketch $\mathbb{S}$, we define the underlying set $|M|$ of M to be the disjoint union of the sets MA, as A ranges over the vertices of $G_\mathbb{S}$; the *cardinality* of M is the cardinality of its underlying set. Then we have

Theorem 2.3.10 *For any sketch $\mathbb{S}$, there exists a regular cardinal κ such that, for any $\mathbb{S}$-model M in* **Set**, *each subset of $|M|$ of cardinality less than κ is contained in (the underlying set of) a submodel of cardinality less than κ.* □

We omit the proof of 2.3.10, for which we refer the reader to [787, Section 3.3], (which also gives information on how κ may be calculated from the size of the various components of $\mathbb{S}$).

Corollary 2.3.11 *For any sketch $\mathbb{S}$, the category $\mathbb{S}$-**Mod**(**Set**) is accessible.*

Proof Let $\mathcal{C}$ be the small category such that models of the linear part of $\mathbb{S}$ (i.e. the sketch obtained from $\mathbb{S}$ by deleting all its cones and cocones) correspond to functors defined on $\mathcal{C}$, as in 2.1.4(*a*). (Note that the objects of $\mathcal{C}$ may be identified with the vertices of $G_\mathbb{S}$.) Then we may consider $\mathbb{S}$-**Mod**(**Set**) as a full subcategory of the functor category $[\mathcal{C}, \mathbf{Set}]$. Let κ be a cardinal such that 2.3.10 holds; we may assume that κ is larger than the cardinality of mor $\mathcal{C}$ (and larger than the cardinality of each cone in $L_\mathbb{S}$). From the latter fact, and the fact that κ-small limits commute with κ-filtered colimits in **Set**, we may conclude that the colimit in $[\mathcal{C}, \mathbf{Set}]$ of any κ-filtered diagram of $\mathbb{S}$-models is again an $\mathbb{S}$-model; in particular, $\mathbb{S}$-**Mod**(**Set**) has κ-filtered colimits.

Let us call a functor $F: \mathcal{C} \to \mathbf{Set}$ κ-small if the cardinality of $|F|$ (i.e. of $\coprod_{A \in \mathrm{ob}\, \mathcal{C}} FA$) is less than κ. Since colimits in $[\mathcal{C}, \mathbf{Set}]$ are computed pointwise, it is easy to see that any κ-small functor is κ-presentable in $[\mathcal{C}, \mathbf{Set}]$. Since the inclusion $\mathbb{S}$-**Mod**(**Set**) $\to [\mathcal{C}, \mathbf{Set}]$ preserves κ-filtered colimits, it follows that if F as above is actually an $\mathbb{S}$-model then it is κ-presentable in $\mathbb{S}$-**Mod**(**Set**). But by 2.3.10, any $\mathbb{S}$-model can be written as the κ-directed union (and hence the κ-filtered colimit) of its κ-small submodels; hence also any κ-presentable object

of $\mathbb{S}$-**Mod**(**Set**) must itself be κ-small, and there are (up to isomorphism) only a set of such models. $\square$

A substantially different proof of 2.3.11, which avoids the use of the Löwenheim–Skolem theorem but is appreciably longer, will be found in [23].

Remark 2.3.12 Let $\mathcal{F}$ be a locally presentable category, and $\mathbb{S}$ a sketch. By 2.3.6, we may identify $\mathcal{F}$ with $\mathbb{T}$-**Mod**(**Set**) for some limit sketch $\mathbb{T}$; it is then not hard to construct a sketch $\mathbb{S} \otimes \mathbb{T}$ such that $(\mathbb{S} \otimes \mathbb{T})$-models in **Set** are 'the same thing as' $\mathbb{S}$-models in $\mathbb{T}$-**Mod**(**Set**). Thus we deduce that, for any such $\mathcal{F}$ and $\mathbb{S}$, the category $\mathbb{S}$-**Mod**($\mathcal{F}$) is accessible; in particular, combining this observation with 2.3.7 and 2.2.7, we deduce that *for any Grothendieck topos $\mathcal{F}$ and any geometric theory $\mathbb{S}$, the category $\mathbb{S}$-**Mod**($\mathcal{F}$) is accessible.* In view of 3.1.13 below, this is equivalent to saying that the 2-category $\mathfrak{BTop}/\mathbf{Set}$ of Grothendieck toposes is 'locally accessible', i.e. all its hom-categories are accessible.

In particular this implies that, *for any two Grothendieck toposes $\mathcal{E}$ and $\mathcal{F}$, there exists a small full subcategory $\mathcal{K}$ of $\mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E})$ such that every object of the latter category is a filtered colimit of objects of $\mathcal{K}$.* Recalling the fact that filtered colimits in $\mathfrak{BTop}/\mathbf{Set}\,(\mathcal{F}, \mathcal{E})$ are computed pointwise at the level of inverse image functors (B3.4.8), this yields an alternative proof of C2.2.11.

For the particular case when $\mathcal{F} = \mathbf{Set}$, a direct proof of the italicized assertion in the second paragraph of 2.3.12, without appealing to the theory of accessible categories, may be found in [504, 7.16].

Example 2.3.13 Before leaving this section, we should give at least one example of a category which is not accessible. We claim that $\mathbf{Set}^{\mathrm{op}}$ *is not accessible.* To see this, let κ be an infinite regular cardinal, let A_κ be a set of cardinality κ, and let J_κ be the poset of subsets of A_κ of cardinality less than κ, ordered by inclusion. It is clear that J_κ is κ-filtered. We have a diagram $D\colon J_\kappa^{\mathrm{op}} \to \mathbf{Set}$ which sends a subset S to $A_\kappa \setminus S$, and an inclusion $S' \subseteq S$ to the inclusion $(A_\kappa \setminus S) \subseteq (A_\kappa \setminus S')$; clearly, the limit of this diagram in **Set** is $\emptyset$. Now if B is any set with more than one element, then any two distinct constant maps $A_\kappa \rightrightarrows B$ remain distinct when restricted to any vertex of the diagram D, but not when restricted to its limit; thus the canonical map $\varinjlim_J(\mathbf{Set}\,(D(S), B)) \to \mathbf{Set}\,(\varprojlim_{J^{\mathrm{op}}}(D(S)), B)$ is not bijective. So B is not κ-presentable in $\mathbf{Set}^{\mathrm{op}}$, for any κ; hence also it cannot be expressed as a colimit in $\mathbf{Set}^{\mathrm{op}}$ of κ-presentable objects.

On the other hand, we know that $\mathbf{Set}^{\mathrm{op}}$ is monadic over **Set** (A2.2.7); so, by the well-known identification of monads on **Set** with infinitary algebraic theories, it can be regarded as the category of models of such a theory. (The theory in question can be presented explicitly as the theory of complete Boolean algebras which are completely distributive, or alternatively as the theory of compact Hausdorff topological Boolean algebras, as we mentioned in A2.2.8.) And any infinitary algebraic theory can be presented by an 'infinite product sketch', as we observed in 2.1.4(c); however, if the theory corresponds (as in this case) to a

monad on **Set** whose functor part is inaccessible (i.e. does not preserve κ-filtered colimits for any κ), then the underlying graph of the sketch will be 'large', i.e. it will have proper classes of vertices and arrows. Thus the inaccessibility of $\mathbf{Set}^{\mathrm{op}}$ is telling us that, in this case at least, the largeness is essential; that is, there is no way of presenting a theory Morita-equivalent to that of completely distributive complete Boolean algebras by means of a sketch based on a small directed graph.

Suggestions for further reading: Adámek & Rosický [23], Ageron [33], Diers [286, 287], Lair [658], Makkai & Paré [787], Makkai & Pitts [788].

D2.4 Properties of model categories

The purpose of this section is to justify a number of remarks which were made in 1.1.7 and elsewhere, to the effect that certain particular theories do not lie in certain of the fragments of first-order logic that we have considered. It is easy to see how one shows that a particular theory is regular (for example): one simply writes down a set of axioms for it, and observes that they are all regular sequents. But how do we show that a theory $\mathbb{T}$ has *no* axiomatization by regular sequents?

The answer, in most cases, is to look at its category of models (in **Set**): if we can show that this category fails to enjoy some categorical property enjoyed by the model categories of all regular theories, then $\mathbb{T}$ cannot be regular. Thus we are interested in finding properties which are in some sense characteristic of the model categories of theories in particular fragments of first-order logic. There are two types of properties in which we shall be interested: 'absolute' properties of $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ as a category (for example, the possession of certain limits or colimits), whose failure for a given $\mathbb{T}$ would show that $\mathbb{T}$ cannot be Morita-equivalent to a regular theory over any signature, and 'relative' properties of $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ as a subcategory of Σ-$\mathbf{Str}(\mathbf{Set})$ (for example, closure under certain limits or colimits), whose failure would show that $\mathbb{T}$ cannot be axiomatized by regular sequents over the particular signature Σ over which it is expressed.

In general, we shall work through the classes of theories in order of increasing complexity, starting with algebraic theories. Before we do so, however, we need to establish some further information about the structure of model categories of cartesian theories. By 2.3.6, we know that such categories are locally finitely presentable (and that every locally finitely presentable category, up to equivalence, occurs in this way). Also, if $\mathbb{T}$ is such a theory, we know that the finitely presentable objects of $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ are exactly (the $\mathbb{T}$-models corresponding to) the representable functors $\mathcal{C}_{\mathbb{T}} \to \mathbf{Set}$, since $\mathcal{C}_{\mathbb{T}}$ has equalizers and so its idempotents split (cf. the remarks after 2.3.3). Let $\{\vec{x}\,.\,\phi\}$ be an object of $\mathcal{C}_{\mathbb{T}}$ (where, by 1.4.4(ii), we may as well assume that ϕ is a finite conjunction of atomic formulae): we shall write $\langle \vec{x} : \phi \rangle$ for the corresponding finitely presentable $\mathbb{T}$-model. This notation is justified by

Lemma 2.4.1 *$\langle \vec{x} : \phi \rangle$ is a representing object for the functor* $\mathbb{T}$-**Mod**(**Set**) $\to$ **Set** *which sends a model M to $[\![\vec{x}.\phi]\!]_M$ (cf.* 1.2.9*).*

Proof By the Yoneda lemma, morphisms $\langle \vec{x} : \phi \rangle \to M$ correspond to elements of $F_M(\{\vec{x}.\phi\})$, where F_M is the cartesian functor $\mathcal{C}_{\mathbb{T}} \to$ **Set** corresponding to M. So the result is immediate from the definition of this functor in the proof of 1.4.7. □

Equivalently, we may think of $\langle \vec{x} : \phi \rangle$ as the $\mathbb{T}$-model M freely generated by a finite string of elements $(a_i \in MA_i \mid 1 \leq i \leq n)$ (where A_i is the sort of x_i), subject to the finite conjunction of conditions corresponding to the formula ϕ. Thus the notion 'finitely presentable', for models of cartesian theories, corresponds to the traditional notion 'finitely presented'. (This is, of course, the origin of the former term.)

Now we turn to algebraic theories.

Lemma 2.4.2 *Let $\mathbb{T}$ be a finitary algebraic theory as in* 1.1.7(a). *Then*

(i) $\mathbb{T}$-**Mod**(**Set**) *is monadic over* **Set**.

(ii) $\mathbb{T}$-**Mod**(**Set**) *is effective regular.*

Proof These are well-known consequences of the characterization of monads on **Set** (see, for example, [726]; cf. also 5.3.1 and 5.3.2 below). □

Similarly, the category of models of a many-sorted algebraic theory as defined in 1.1.7(b) is monadic over some power of **Set**, and hence effective regular. Thus, for example, the theory of torsion-free abelian groups, described in 1.1.7(d), cannot be Morita-equivalent to an algebraic theory, for we saw in Section A1.3 that its category of models is (regular but) not effective. The latter state of affairs is typical of Horn theories over signatures having no relation symbols (other than equality), as we shall see in 2.4.4 below. But first we need

Lemma 2.4.3 *Let $\mathbb{T}$ be a Horn theory over a signature Σ. Then*

(i) *An arbitrary product (in* Σ-**Str**(**Set**)*) of $\mathbb{T}$-models is a $\mathbb{T}$-model.*

(ii) *Any substructure of a $\mathbb{T}$-model is a $\mathbb{T}$-model.*

Proof (i) It is easy to verify that, for any family $(M_i \mid i \in I)$ of Σ-structures and any Horn formula-in-context $\vec{x}.\phi$, the interpretation of $\vec{x}.\phi$ in the product $\prod_{i \in I} M_i$ is simply the product of the $[\![\vec{x}.\phi]\!]_{M_i}$. Hence any Horn sequent satisfied in each M_i is satisfied in the product.

(ii) Similarly, if N is a substructure of M, then as we remarked in 1.2.10(c) the interpretation in N of any quantifier-free formula-in-context (in particular, of any Horn formula-in-context) $\vec{x}.\phi$ is simply the intersection of $[\![\vec{x}.\phi]\!]_M$ with $N(A_1, \ldots, A_n)$. So N satisfies any Horn sequent satisfied by M. □

Part (i) of 2.4.3 remains true for arbitrary regular theories, by essentially the same proof, but part (ii) does not (even for cartesian theories): for the theory of categories as described in 1.1.7(e), it is easily seen that there are substructures of models which are not models. Hence the theory of categories cannot be axiomatized by Horn sequents over the particular signature considered in 1.1.7(e). After 2.4.6 below, we shall be able to prove that it cannot be a Horn theory over any signature.

Corollary 2.4.4 *Let* $\mathbb{T}$ *be a Horn theory over a signature* Σ *such that* Σ-Rel *is empty. Then*

(i) $\mathbb{T}$-**Mod**(**Set**) *is reflective in* Σ-**Str**(**Set**), *the unit of the reflection being regular epic.*

(ii) $\mathbb{T}$-**Mod**(**Set**) *is a regular category.*

Proof (i) Given a Σ-structure M and a family of congruences $(R_i \mid i \in I)$ on M (that is, 'sortwise' equivalence relations on M which are sub-Σ-structures of $M \times M$) such that each of the quotients M/R_i is a $\mathbb{T}$-model, the same is true of $M/(\bigcap_{i\in I} R_i)$ (since it maps injectively into the product of the M/R_i). Hence there is a unique smallest congruence R on M such that M/R is a $\mathbb{T}$-model, namely the intersection of all such congruences. It is then clear that the assignment $M \mapsto M/R$ defines a left adjoint to the inclusion $\mathbb{T}$-**Mod**(**Set**) $\to$ Σ-**Str**(**Set**), the unit of the adjunction being the quotient map.

(ii) It follows easily from (i) that regular epimorphisms in $\mathbb{T}$-**Mod**(**Set**) are preserved by the inclusion into Σ-**Str**(**Set**), and hence that they are exactly the (sortwise) surjective homomorphisms of $\mathbb{T}$-models. Hence they are stable under pullback, and $\mathbb{T}$-**Mod**(**Set**) is regular. □

The category **Poset** of partially ordered sets is not regular: if **n** denotes an n-element totally ordered set, then there is a regular epimorphism $\mathbf{2} \amalg \mathbf{2} \to \mathbf{3}$ in **Poset** (obtained by identifying the top element of one copy of **2** with the bottom element of the other), whose pullback along the third injection $\mathbf{2} \rightarrowtail \mathbf{3}$ is (monic, and hence) not a cover. Since we saw in 1.1.7(c) that posets are the models of a Horn theory, this shows that the restriction on Σ in the statement of 2.4.4 cannot be omitted – and also that there can be no analogue of 1.4.9 for Horn theories, as we claimed in Section D1.4.

However, Horn theories over signatures without function symbols also have their limitations. We note that **Poset**, although not a regular category, resembles a coherent category to the extent of having a strict initial object (cf. A1.4.1), and this state of affairs is again typical:

Lemma 2.4.5 *Let* Σ *be a signature such that* Σ-Fun *is empty, and* $\mathbb{T}$ *a Horn theory over* Σ. *Then* $\mathbb{T}$-**Mod**(**Set**) *has a strict initial object.*

Proof The initial object M of $\mathbb{T}$-**Mod**(**Set**) (the free $\mathbb{T}$-model on no generators) may be constructed as follows. MA is empty for each sort A, which forces

MR to be empty for each relation symbol R of arity ≥ 1. For nullary relations R, we also take $MR = \emptyset$ unless we are forced to do otherwise by the fact that $(\top \vdash_{[]} R)$ is provable in $\mathbb{T}$. It is straightforward to verify that this Σ-structure is a $\mathbb{T}$-model, and that it is strict initial in $\mathbb{T}$-**Mod**(**Set**). □

Since the initial object of **TF** is also terminal, we see that the theory of 1.1.7(d) cannot be Morita-equivalent to a Horn theory over a signature without function symbols. (However, 2.4.5 does not characterize the model categories of Horn theories over relational signatures: the same argument would work if Σ contained function symbols of arity ≥ 1, provided there were no constants.)

The property which distinguishes Horn theories (over arbitrary signatures) from arbitrary cartesian theories is the presence of 'enough projectives' in their categories of models.

Lemma 2.4.6 *Let $\mathbb{T}$ be a Horn theory over a signature Σ. Then*

(i) *Any cover $h\colon M \twoheadrightarrow N$ in $\mathbb{T}$-**Mod**(**Set**) is a sortwise surjection, i.e., for each sort A of Σ, $h_A\colon MA \to NA$ is surjective.*

(ii) *$\mathbb{T}$-**Mod**(**Set**) has a separating set of objects which are projective with respect to covers.*

Proof (i) Given a homomorphism $h\colon M \to N$ of $\mathbb{T}$-models, it is clear that the sortwise image of h can be made into a substructure of N, and hence a $\mathbb{T}$-model by 2.4.3(ii). So if h is a cover this image must be the whole of N, i.e. h is sortwise surjective.

(ii) For each sort A of Σ, the functor $\mathbb{T}$-**Mod**(**Set**) $\to$ **Set** which sends M to MA is representable, by 2.4.1. Part (i) implies that the representing objects $\langle x : \top \rangle$ are projective with respect to covers; and since the family of all such functors, as A ranges over the sorts of Σ, is clearly faithful, they form a separating set. □

If the signature Σ has no relation symbols, we may strengthen 'separating' to 'generating' in 2.4.6(ii), by essentially the same proof.

Example 2.4.7 We may now prove, as promised above, that the theory of categories cannot be a Horn theory over any signature: for covers (in particular, regular epimorphisms) in **Cat** do not have to be surjective on morphisms, from which one may easily deduce that the only cover-projectives in **Cat** are discrete categories. And since it is possible for two functors to agree on objects but not on morphisms, these do not form a separating family for **Cat**.

For any cartesian theory $\mathbb{T}$, the category $\mathbb{T}$-**Mod**(**Set**) is closed in Σ-**Str**(**Set**) not only under products, as we observed after 2.4.3, but also under equalizers (by a similarly straightforward proof) and hence under arbitrary limits. (Of course, we knew already from 2.3.6 that $\mathbb{T}$-**Mod**(**Set**) has all small limits – and colimits.) Thus the theory of divisible abelian groups, mentioned in 1.1.7(f), is not expressible as a cartesian theory over the signature for groups, since the

category **Div** of divisible abelian groups is not closed under equalizers in **Ab**. In this case **Div** is complete and cocomplete, since it is coreflective in **Ab**; but it is not locally finitely presentable – it can be shown that the only finitely presentable object of **Div** is the trivial group – and so it cannot be the category of models of a cartesian theory over any signature. Similar remarks apply to the regular theory considered in 2.3.9.

For regular theories, we still have closure under products, as mentioned after 2.4.3, and hence in particular the category of models has products, though it may lack equalizers. (A simple counterexample is the theory $\mathbb{O}_1$ of inhabited objects (cf. C5.2.8(c)), i.e. the theory over the signature with one sort and no primitive symbols except equality, whose only axiom is $(\top \vdash_{[]} (\exists x)\top)$: its category of models is the category of nonempty sets, and the two morphisms $1 \rightrightarrows 2$ in this category have no equalizer.) Thus the theory of fields, considered in 1.1.7(h), cannot be Morita-equivalent to a regular theory over any signature, since its category of models lacks a terminal object.

Remark 2.4.8 As previously, we shall not give detailed consideration to the class of disjunctive theories, introduced in 1.3.6; but it seems worth mentioning that *if* $\mathbb{T}$ *is a disjunctive theory over a signature* Σ, *then* $\mathbb{T}$-**Mod**(**Set**) *is closed under arbitrary connected limits in* Σ-**Str**(**Set**), and in particular it has connected limits. The proof is, as usual, a straightforward induction to show that the interpretations of disjunctive formulae are 'preserved' under connected limits. Using this result we may show that, for example, the theory of local rings (1.1.7(g)) cannot be Morita-equivalent to a disjunctive theory: for the diagram

$$\begin{tikzcd} & \mathbb{Z}_{(2)} \arrow[d, tail] \\ \mathbb{Z}_{(3)} \arrow[r, tail] & \mathbb{Q} \end{tikzcd}$$

of local rings (where $\mathbb{Z}_{(p)}$ denotes the ring of rationals whose denominators are prime to p) has no pullback – indeed, there is no cone over it in the category of local rings, since in any local ring either -2 or $+3$ is invertible.

We saw in the last section that the category of models of a geometric theory, although it is always accessible, need not be finitely accessible. Nevertheless, it does always possess filtered (that is, ω-filtered) colimits. To prove this, we need another closure result, similar to those already established:

Lemma 2.4.9 *Let* $\mathbb{T}$ *be a geometric theory over a signature* Σ. *Then* $\mathbb{T}$-**Mod**(**Set**) *is closed under filtered colimits in* Σ-**Str**(**Set**).

Proof Let $(M_i \mid i \in I)$ be the vertices of a filtered diagram of Σ-structures, with colimit N. Since the structure used in interpreting geometric formulae is

all derived from finite limits and arbitrary colimits, and these commute with filtered colimits in **Set**, it is straightforward to verify that, for any geometric formula-in-context $\vec{x}.\phi$, the interpretation $[\![\vec{x}.\phi]\!]_N$ is the colimit of the $[\![\vec{x}.\phi]\!]_{M_i}$. Hence N satisfies any geometric sequent which is satisfied by all the M_i. $\square$

In view of 3.1.13 below, 2.4.9 is really no more than a restatement of (a particular case of) Lemma B3.4.8.

Example 2.4.10 We may now show that the theory $\mathbb{T}$ of posets in which every element lies below a maximal element, described in 1.1.7(i), is not Morita-equivalent to a coherent theory (or even a geometric theory). For every natural number n, the n-element totally ordered set $\mathbf{n} = \{0, 1, \ldots, n-1\}$ is a model of $\mathbb{T}$, and the inclusions $\mathbf{n} \rightarrowtail \mathbf{m}$ for $n < m$ are $\mathbb{T}$-model homomorphisms, which form a filtered diagram in $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$. The colimit of this diagram in Σ-$\mathbf{Str}(\mathbf{Set})$ is the ordered set of natural numbers, which is not a model of $\mathbb{T}$; so it is immediate that $\mathbb{T}$ cannot be axiomatized by geometric sequents over the signature Σ. But in fact it is not hard to see that this diagram has no colimit in $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ – although we can construct a cone under it by adjoining a top element to the set of natural numbers, this is not a colimit because maximal elements do not have to be preserved by $\mathbb{T}$-model homomorphisms – and so $\mathbb{T}$ cannot be equivalent to a geometric theory over any signature. A straightforward generalization of this argument, using successor ordinals regarded as $\mathbb{T}$-models, shows that $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ does not have κ-filtered colimits for any κ; hence it is not accessible, and so $\mathbb{T}$ is not 'Morita-equivalent' to any sketch, let alone a finitary one.

On the other hand, if $\mathbb{T}$ is any (finitary) first-order theory, then it follows from 1.5.13 and 2.4.9 that $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})_e$ does have filtered colimits. And M. Richter [1017] has shown that, for any such $\mathbb{T}$, any filtered colimits which exist in $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$ must be preserved by the forgetful functor to Σ-$\mathbf{Str}(\mathbf{Set})$: that is, if the colimit in Σ-$\mathbf{Str}(\mathbf{Set})$ of a filtered diagram of $\mathbb{T}$-models fails to be a $\mathbb{T}$-model, then the diagram cannot have a colimit in $\mathbb{T}$-$\mathbf{Mod}(\mathbf{Set})$. (See also [1059].)

In a similar vein, the theory of metric spaces cannot be axiomatized by geometric sequents over the signature of 1.1.7(l), since its models fail to be closed under filtered colimits (take a directed diagram of two-point spaces indexed by the natural numbers, such that the distance between the two points in the nth space is 2^{-n}). However, the category of metric spaces (and non-expansive mappings) does have filtered colimits, since it is reflective in the category of pseudometric spaces, and we observed in 1.1.7(l) that the latter are the models of a geometric theory.

To distinguish finitary first-order theories from infinitary ones, the best tool we have available is the notion of ultraproduct, which in most cases yields a 'relative' rather than an 'absolute' criterion. Given a family of Σ-structures $(A_i \mid i \in I)$ and a filter Φ on (the power-set of) I, we define the *filterproduct* $\prod_\Phi A_i$

to be the quotient of the product $\prod_{i \in I} A_i$ obtained by identifying two elements $(a_i \mid i \in I)$ and $(a'_i \mid i \in I)$ iff the set $\{i \in I \mid a_i = a'_i\}$ lies in Φ. (This definition requires some modification if (some of the sorts of) some of the A_i are allowed to be empty, but we shall not bother about this.) It is clear that $\prod_\Phi A_i$ can be given a Σ-structure in a canonical way. If Φ is an ultrafilter (in which case we call $\prod_\Phi A_i$ an *ultraproduct*), then an easy induction shows that for any first-order formula-in-context $\vec{x}\,.\,\phi$, an n-tuple of elements $b_1, \ldots, b_n$ of $\prod_\Phi A_i$ lies in the interpretation of $\vec{x}\,.\,\phi$ iff, given any choice of representatives $(a_{ji} \mid i \in I)$ for each b_j, the set

$$\{i \in I \mid (a_{1i}, \ldots, a_{ni}) \in [\![\vec{x}\,.\,\phi]\!]_{A_i}\}$$

is in Φ. We thus obtain the result known to model-theorists as Łoś's Theorem:

Lemma 2.4.11 *For any (finitary) first-order theory $\mathbb{T}$ over a signature Σ, the class of $\mathbb{T}$-models is closed under ultraproducts in the class of Σ-structures.* □

The reason why this does not lead to an absolute condition characterizing model categories of first-order theories is that the ultraproduct construction cannot be characterized in abstract categorical terms; it is a piece of extra structure carried by the model category. The 'ultracategories' studied by M. Makkai [778, 780] are in effect categories equipped with extra structure of this type.

Examples 2.4.12 (a) Nevertheless, we may now show (as promised in 2.1.4(f)) that the theory of connected directed graphs, although it can be axiomatized by a finitary sketch, cannot be axiomatized by any finitary first-order theory over the signature for directed graphs. For each natural number $n > 0$, let G_n denote the graph with n vertices $1, 2, \ldots, n$ and a chain of $(n-1)$ arrows joining them; then G_n is connected for each n. But if Φ is any non-principal ultrafilter on the set $\mathbb{N}$ of natural numbers, then $\prod_\Phi G_n$ is not connected, since the vertices which are the equivalence classes of the sequences $(1 \mid n \in \mathbb{N})$ and $(n \mid n \in \mathbb{N})$ cannot be joined by any finite chain.

(b) In the same way, we may construct ultraproducts of torsion groups which have elements of infinite order; so the geometric theory of torsion groups, described in 1.1.7(j), cannot be axiomatized by finitary sequents over the signature for groups.

(c) The theory $\mathbb{K}$ of K-finite objects, axiomatized in 1.1.7(k), cannot be Morita-equivalent to a finitary theory over any signature. To see this, note that if $(A_i \mid i \in I)$ is any family of models of a finitary theory, and Φ is an ultrafilter on I, then the ultraproduct $\prod_\Phi \mathrm{Aut}(A_i)$ of the automorphism groups of the A_i embeds in the obvious way as a subgroup of the automorphism group of $\prod_\Phi A_i$. But $\mathbb{K}$ has models (in **Set**) with arbitrarily large finite automorphism groups; so a nontrivial ultraproduct of these would have to have an infinite automorphism group.

We note that all three of the theories i
axiomatized by finitary sketches, using 2.
for the theory of 2.4.12(a).) This emphasiz
have met at several points in this chapter
version of finitariness, as expressed in the
well with the more traditional logical vers
(first-order) theory.

Suggestions for further reading: Héb
660, 661], Makkai [778, 779, 780, 783], Ricl

ı 2.4.12 are σ-coherent, so they *can* be
2.9. (Of course, we already knew this
es once again a phenomenon which we
: namely, that the 'category-theoretic'
notion of finitary sketch, does not fit
ion expressed in the notion of finitary

ert [437, 438], Keane [575], Lair [659,
ıter [1017], Rosický [1053, 1058, 1059].

D3

CLASSIFYING TOPOSES

D3.1 Classifying toposes via syntactic sites

We saw in Section D1.4 that, for each cartesian (resp. regular, coherent, geometric) theory $\mathbb{T}$, there is a category $\mathcal{C}_{\mathbb{T}}$ of the appropriate kind containing a model of $\mathbb{T}$ which is generic, in the sense defined there. Since we are mainly interested in models which live in toposes, it would be convenient to be able to construct a topos containing a model of a given theory which is generic amongst models in toposes. There are various methods for doing this. One, which was sketched in Section B4.2, relies on the 2-categorical structure of (Grothendieck) toposes, in order to build up (the topos containing) the generic model by an inductive process. Here we shall give a more 'pedestrian', but also more explicit, method, which has the advantage that the explicit description of the generic model which we obtain allows us to establish results about models of a theory by first proving them in the generic case, and then transferring them along inverse image functors. (We shall see several examples of this technique in the subsequent sections of this chapter, as well as in Chapters F1 and F4.)

The essence of this construction is to use the syntactic category $\mathcal{C}_{\mathbb{T}}$ itself as the underlying category of a site (the *syntactic site* of the theory), by imposing a suitable coverage J on it in such a way that cartesian (resp. regular, coherent, geometric) functors defined on $\mathcal{C}_{\mathbb{T}}$ are equivalent to geometric morphisms (over **Set**) into $\mathbf{Sh}(\mathcal{C}_{\mathbb{T}}, J)$. (As in Section B4.2, we shall denote this topos by $\mathbf{Set}[\mathbb{T}]$, and it will often be fruitful to think of it as the category of all 'geometric constructs', in the sense defined there, in the theory $\mathbb{T}$.) There are (at least) four different versions of the construction, corresponding to the four different constructions of $\mathcal{C}_{\mathbb{T}}$ appropriate to the different fragments of logic under consideration. As in Section D1.4, we shall take them in order of increasing logical complexity. (Of course, every cartesian theory is regular, every regular theory is coherent, and every coherent theory is geometric, so we don't really need four different constructions – the one for geometric theories would suffice. However, the emphasis in this section is not on the mere existence of classifying toposes, but rather on obtaining the most explicit description possible of the classifying topos, and in particular of the generic model, of a given theory.)

For cartesian theories, in fact, no coverage is needed. We saw in 1.4.7 that, for any cartesian theory $\mathbb{T}$, we have a natural equivalence

$$\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E}) \simeq \mathfrak{Cart}\,(\mathcal{C}_{\mathbb{T}}, \mathcal{E})$$

for any cartesian category $\mathcal{E}$. But if $\mathcal{E}$ is a **Set**-topos, then we also have a natural equivalence

$$\mathfrak{Cart}\,(\mathcal{C}_{\mathbb{T}}, \mathcal{E}) \simeq \mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, [(\mathcal{C}_{\mathbb{T}})^{\mathrm{op}}, \mathbf{Set}])$$

by Diaconescu's theorem (B3.2.7). So, on combining these two equivalences, we deduce

Theorem 3.1.1 *Let $\mathbb{T}$ be a cartesian theory. Then the topos $[(\mathcal{C}_{\mathbb{T}})^{\mathrm{op}}, \mathbf{Set}]$ contains a $\mathbb{T}$-model $G_{\mathbb{T}}$ such that, for any* **Set***-topos $\mathcal{E}$, the functor*

$$\mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, [(\mathcal{C}_{\mathbb{T}})^{\mathrm{op}}, \mathbf{Set}]) \longrightarrow \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$$

which sends a geometric morphism f to $f^*(G_{\mathbb{T}})$ is one half of an equivalence of categories. □

Since the second of the two equivalences above is obtained (in one direction) by composing with the Yoneda embedding $\mathcal{C}_{\mathbb{T}} \to [(\mathcal{C}_{\mathbb{T}})^{\mathrm{op}}, \mathbf{Set}]$, the generic model $G_{\mathbb{T}}$ of 3.1.1 is simply the image under this embedding of the generic model $M_{\mathbb{T}}$ of 1.4.7. But in 2.4.1 we saw that the representable functors on $\mathcal{C}_{\mathbb{T}}$ may be identified with finitely presented $\mathbb{T}$-models in **Set**. Thus we obtain

Corollary 3.1.2 *For a cartesian theory $\mathbb{T}$, the classifying topos* $\mathbf{Set}[\mathbb{T}]$ *may be identified with the functor category* $[\mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})_\omega, \mathbf{Set}]$, *where* $\mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})_\omega$ *denotes the category of finitely presented $\mathbb{T}$-models; and $G_{\mathbb{T}}$ may be identified with the inclusion functor* $\mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})_\omega \to \mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})$, *regarded as in* 1.2.14(i) *as a $\mathbb{T}$-model in the functor category – that is, for each sort A of the signature of $\mathbb{T}$, $G_{\mathbb{T}}A$ is the functor which sends a finitely presented $\mathbb{T}$-model M to MA, and similarly for function symbols and relation symbols.* □

This explicit description of $G_{\mathbb{T}}$ will be of use on many occasions later on. Notice, by the way, that the syntactic category $\mathcal{C}_{\mathbb{T}}$ is recoverable up to equivalence from the classifying topos $\mathbf{Set}[\mathbb{T}]$, since it is dual to the full subcategory of finitely presentable objects of $\mathfrak{Top}/\mathbf{Set}\,(\mathbf{Set}, \mathbf{Set}[\mathbb{T}])$. That is, *the classifying topos of a cartesian theory determines the theory up to Morita equivalence.*

As an example of the applications of 3.1.2, we give

Corollary 3.1.3 *Let $\mathbb{T}$ be a finitary algebraic theory in the sense of* 1.1.7(*a*). *Then the generic $\mathbb{T}$-model $G_{\mathbb{T}}$ is* functionally complete *in the sense that, for any natural number n, each morphism $(G_{\mathbb{T}})^n \to G_{\mathbb{T}}$ in* $\mathbf{Set}[\mathbb{T}]$ *is the interpretation in $G_{\mathbb{T}}$ of some n-ary derived operation of $\mathbb{T}$ (that is, of some term-in-context $x_1, \ldots, x_n . t$). Moreover, two such terms-in-context define the same morphism $(G_{\mathbb{T}})^n \to G_{\mathbb{T}}$ iff the sequent expressing their equality is provable in $\mathbb{T}$.*

Proof Since (the underlying object of) $G_{\mathbb{T}}$ is the forgetful functor from finitely-presented $\mathbb{T}$-models to **Set**, it is representable by the free $\mathbb{T}$-model $F(1)$ on one generator; similarly, $(G_{\mathbb{T}})^n$ is representable by the free model $F(n)$ on n generators. So the Yoneda lemma tells us that morphisms $(G_{\mathbb{T}})^n \to G_{\mathbb{T}}$ correspond bijectively to morphisms $F(1) \to F(n)$, or equivalently to elements of the underlying set of $F(n)$; but the latter are exactly the provable-equality-classes of n-ary derived operations. □

For example, if $\mathbb{T}$ is the theory of commutative rings (with 1), then the morphisms $(G_{\mathbb{T}})^n \to G_{\mathbb{T}}$ in $\mathbf{Set}[\mathbb{T}]$ are exactly the polynomial functions of n variables (with integer coefficients). Although we stated 3.1.3 only for single-sorted theories, it remains valid (with the same proof – but a more complicated statement) for many-sorted algebraic theories as defined in 1.1.7(b).

We turn now to regular theories. We recall that if $\mathcal{C}$ is a small regular category, the *regular coverage* R on $\mathcal{C}$ has as covering sieves all sieves which contain a cover in the sense defined in Section A1.3. In A2.1.11(a), we saw that the representable functors $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ are all sheaves for this coverage, so that the Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ factors through $\mathbf{Sh}(\mathcal{C}, R)$; and the resulting full embedding $\mathcal{C} \to \mathbf{Sh}(\mathcal{C}, R)$ is (not only cartesian but) a regular functor. Further, for any Grothendieck topos $\mathcal{E}$, a cartesian functor $F: \mathcal{C} \to \mathcal{E}$ is cover-preserving iff it maps each sieve generated by a single cover in $\mathcal{C}$ to an epimorphic family, iff it maps covers in $\mathcal{C}$ to epimorphisms in $\mathcal{E}$, iff it is a regular functor. So by C2.3.9 we have an equivalence

$$\mathfrak{Reg}\,(\mathcal{C}, \mathcal{E}) \simeq \mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, \mathbf{Sh}(\mathcal{C}, R))$$

for any such topos $\mathcal{E}$. Applying the above to the syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ of a regular theory $\mathbb{T}$, we immediately obtain

Theorem 3.1.4 *For any regular theory $\mathbb{T}$, the topos $\mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}, R)$ contains a $\mathbb{T}$-model $G_{\mathbb{T}}$ such that, for any Grothendieck topos $\mathcal{E}$, the functor*

$$\mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, \mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}, R)) \longrightarrow \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$$

sending f to $f^*(G_{\mathbb{T}})$ is one half of an equivalence of categories. Moreover, the regular sequents satisfied by $G_{\mathbb{T}}$ are precisely those provable in $\mathbb{T}$.

Proof Combine the regular version of 1.4.7 with the remarks above. The $\mathbb{T}$-model $G_{\mathbb{T}}$ is of course the image of $M_{\mathbb{T}}$ under the embedding $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}} \to \mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}})$; the last assertion of the theorem follows from the fact that this embedding is full and therefore conservative. □

Remark 3.1.5 Since any topos $\mathcal{E}$ is effective as a regular category, any regular functor $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}} \to \mathcal{E}$ factors through the effectivization $\mathcal{E}_{\mathbb{T}}$ of $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ (cf. 1.4.12(i)). So we could have used $\mathcal{E}_{\mathbb{T}}$, with its regular coverage, as a site of definition for the classifying topos of $\mathbb{T}$. This fact, combined with the existence of small regular

categories which are not effective, indicates that we cannot hope to recover the syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ up to equivalence from the classifying topos $\mathbf{Set}[\mathbb{T}]$, as we could in the case of cartesian theories. However, we shall see in 3.3.10 below that we can recover $\mathcal{E}_{\mathbb{T}}$ up to equivalence from $\mathbf{Set}[\mathbb{T}]$ – that is, we can recover the theory $\mathbb{T}$ up to Morita equivalence as defined after 1.4.12.

Remark 3.1.6 We noted after 1.4.11 that if $\mathbb{T}$ is a cartesian theory then its regular syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$ is equivalent to the regularization of its cartesian syntactic category $\mathcal{C}_{\mathbb{T}}$. But in C2.2.19 we noted that, for any small cartesian category $\mathcal{C}$, the topos of sheaves on $\mathbf{Reg}(\mathcal{C})$ for the regular coverage is equivalent to $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. Thus we have a direct proof that 3.1.1 and 3.1.4 produce equivalent constructions of $\mathbf{Set}[\mathbb{T}]$.

There is an alternative construction of $\mathbf{Set}[\mathbb{T}]$, less 'canonical' but often more convenient to use, which in some sense stands in the same relation to 3.1.4 as 3.1.2 does to 3.1.1. Given a regular theory $\mathbb{T}$, we can always find a cartesian theory $\mathbb{T}_0$ over the same signature, all of whose axioms are provable in $\mathbb{T}$. (We could take $\mathbb{T}_0$ to be the empty theory in this signature, or the theory consisting of all the Horn sequents provable in $\mathbb{T}$; but $\mathbb{T}_0$ need not be a Horn theory – for example, if $\mathbb{T}$ were the theory of strongly connected categories (categories in which, given any two objects, there exists a morphism from the first to the second; cf. A4.6.9), it would be natural to take $\mathbb{T}_0$ to be the theory of categories.) Then, by 1.3.10(ii), we may write the axioms of $\mathbb{T}$ (other than those which occur as, or are derivable from, axioms of $\mathbb{T}_0$) in the form $(\phi \vdash_{\vec{x}} (\exists \vec{y})\psi)$, where $\vec{x}\,.\,\phi$ and $\vec{x}, \vec{y}\,.\,\psi$ are cartesian formulae-in-context relative to $\mathbb{T}_0$ (indeed, they could be Horn formulae), and the sequent $(\psi \vdash_{\vec{x},\vec{y}} \phi)$ is provable in $\mathbb{T}_0$. This latter condition means that, in the syntactic category $\mathcal{C}_{\mathbb{T}_0}$ of $\mathbb{T}_0$, we have a morphism

$$\{\vec{x}, \vec{y}\,.\,\psi\} \xrightarrow{[(\psi \wedge (\vec{x} = \vec{x}'))]} \{\vec{x}'\,.\,\phi[\vec{x}'/\vec{x}]\}\ ,$$

and a $\mathbb{T}_0$-model M in a regular category $\mathcal{C}$ is a $\mathbb{T}$-model iff, for each axiom of $\mathbb{T}$, the functor F_M corresponding to M via 1.4.7 sends this morphism to a cover in $\mathcal{C}$. Thus we obtain

Proposition 3.1.7 *With the above notation, the classifying topos* $\mathbf{Set}[\mathbb{T}]$ *may be identified with* $\mathbf{Sh}(\mathcal{C}, J_{\mathbb{T}})$ *where* $\mathcal{C}$ *is the syntactic category of* $\mathbb{T}_0$ *(equivalently, the opposite of the category of finitely presented* $\mathbb{T}_0$*-models in* $\mathbf{Set}$*), and* $J_{\mathbb{T}}$ *is the smallest coverage on this category for which the sieve generated by the single morphism displayed above is* $J_{\mathbb{T}}$*-covering for each axiom of* $\mathbb{T}$ *not occurring among (or derivable from) the axioms of* $\mathbb{T}_0$. □

The usefulness of this result is greatly dependent on the extent to which we can give an explicit description of the coverage $J_{\mathbb{T}}$, and in particular on whether we can prove it to be subcanonical (i.e. whether we can prove that all representable functors are sheaves for it). This is because the generic $\mathbb{T}$-model $G_{\mathbb{T}}$

is simply the associated $J_{\mathbb{T}}$-sheaf of the generic $\mathbb{T}_0$-model; but since the latter consists of representable functors by 3.1.2, it will already be a sheaf if $J_{\mathbb{T}}$ is subcanonical, and therefore we shall have an explicit description of $G_{\mathbb{T}}$ at no extra cost. Clearly, the question whether $J_{\mathbb{T}}$ is subcanonical will depend (not just on $\mathbb{T}$ but) on the choice of the theory $\mathbb{T}_0$: the stronger we make this theory, the fewer the covering families we shall have to put into $J_{\mathbb{T}}$, and the more likely it becomes that we shall obtain a standard site.

Example 3.1.8 If $\mathbb{T}$ is the theory of strongly connected categories, mentioned above, and we take $\mathbb{T}_0$ to be the theory of categories, then 3.1.7 tells us to find the smallest coverage on the opposite of the category $\mathbf{Cat}_\omega$ of finitely presented categories for which the inclusion

$$2 \xrightarrow{\ i\ } \mathbf{2}$$

where 2 is the two-object discrete category and $\mathbf{2}$ is the two-element totally ordered set, generates a covering sieve on 2. By the results of Section C2.1, a functor $F\colon \mathbf{Cat}_\omega \to \mathbf{Set}$ will be a sheaf for this coverage iff it satisfies the sheaf axiom for (the sieves generated by) the pushouts of i along arbitrary morphisms $2 \to \mathcal{C}$, i.e. for the sieves generated by morphisms of the form $\mathcal{C} \to \mathcal{C}[f]$, where $\mathcal{C}[f]$ denotes the category obtained by freely adjoining a morphism f between two given objects of $\mathcal{C}$. But it is easy to verify that all such morphisms are regular monomorphisms in $\mathbf{Cat}$ (i.e. they are the equalizers of their cokernel-pairs), and hence in $\mathbf{Cat}_\omega$; so the representable functors are indeed sheaves. Thus we may describe the generic strongly connected category explicitly: it is simply given by the same data as the generic category, but regarded as an internal category in the sheaf topos $\mathbf{Sh}(\mathbf{Cat}_\omega^{\mathrm{op}}, J_{\mathbb{T}})$ rather than in the functor category $[\mathbf{Cat}_\omega, \mathbf{Set}]$.

On the other hand, if we took $\mathbb{T}_0$ to be the 'Horn part' of the theory of categories (i.e. the theory obtained from that of categories by omitting the third displayed axiom of 1.1.7(e); we may think of models of this theory as 'partial categories', in which some composites may be undefined), then our coverage on $\mathcal{C}^{\mathrm{op}}$ (where $\mathcal{C}$ is the category of finitely presented partial categories) would have to contain the sieve generated by the inclusion

$$\begin{tikzcd}
\bullet \arrow[r, "y"] & \bullet \arrow[dd, "x"] & & \bullet \arrow[r, "y"] \arrow[ddr, "z"'] & \bullet \arrow[dd, "x"] \\
& & {} \arrow[r] & {} & \\
& \bullet & & & \bullet
\end{tikzcd}$$

where, in the right-hand partial category, z is specified as being the composite of x and y. And this morphism is a non-invertible epimorphism in $\mathcal{C}$, so it cannot be a regular monomorphism; hence not all representable functors are sheaves for this coverage.

We leave it to the reader to verify that the classifying topos for the theory of divisible abelian groups (1.1.7(f)) may be constructed by means of a subcanonical coverage on the opposite of the category $\mathbf{Ab}_\omega$ of finitely-presented abelian groups.

For coherent theories, the arguments are not very different from those for regular theories. If $\mathcal{C}$ is a small coherent category, we defined the *coherent coverage* P on $\mathcal{C}$ in A2.1.11(b) to consist of all finite families $(f_i \colon A_i \to A \mid 1 \leq i \leq n)$ such that the union of the images of the f_i is the whole of A. We observed that this coverage is subcanonical, so that the Yoneda embedding becomes a full embedding $\mathcal{C} \to \mathbf{Sh}(\mathcal{C}, P)$, and that this functor is coherent. And it is easy to verify that a cartesian functor from $\mathcal{C}$ to a topos sends finite families as above to epimorphic families iff it preserves covers and finite unions of subobjects, i.e. iff it is a coherent functor. So we obtain

Theorem 3.1.9 *For any coherent theory $\mathbb{T}$, the topos $\mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}, P)$ contains a $\mathbb{T}$-model $G_{\mathbb{T}}$ such that, for any Grothendieck topos $\mathcal{E}$, the functor*

$$\mathfrak{Top}/\mathbf{Set}\,(\mathcal{E}, \mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}, P)) \longrightarrow \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$$

sending f to $f^*(G_{\mathbb{T}})$ is one half of an equivalence of categories. Moreover, the coherent sequents satisfied by $G_{\mathbb{T}}$ are precisely those provable in $\mathbb{T}$. □

As in 3.1.5, we could have replaced $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ by its pretopos completion $\mathcal{P}_{\mathbb{T}}$ (cf 1.4.12(ii)) in the statement of the above theorem; and in 3.3.8 below we shall see that $\mathcal{P}_{\mathbb{T}}$, though not $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$, is recoverable up to equivalence from the topos $\mathbf{Set}[\mathbb{T}]$.

In the opposite direction, we can make the underlying category of the site smaller, as in 3.1.7, by regarding $\mathbb{T}$ as a 'quotient' of a cartesian theory $\mathbb{T}_0$ obtained by adding new axioms of the form

$$\left(\phi \vdash_{\vec{x}} \bigvee_{i=1}^{n} (\exists \vec{y_i})\psi_i\right)$$

where $\vec{x}\,.\,\phi$ and $\vec{x}, \vec{y_i}\,.\,\psi_i$ are cartesian formulae-in-context relative to $\mathbb{T}_0$ and the sequents $(\psi_i \vdash_{\vec{x},\vec{y_i}} \phi)$ are provable in $\mathbb{T}_0$ (cf. 1.3.10(iii)). For each such axiom, we then have a family of morphisms

$$\{\vec{x}, \vec{y_i}\,.\,\psi_i\} \xrightarrow{[(\psi_i \wedge (\vec{x} = \vec{x'}))]} \{\vec{x'}\,.\,\phi[\vec{x'}/\vec{x}]\}$$

$(1 \leq i \leq n)$ in the syntactic category $\mathcal{C}_{\mathbb{T}_0}$, and a $\mathbb{T}_0$-model is a $\mathbb{T}$-model iff the cartesian functor on $\mathcal{C}_{\mathbb{T}_0}$ corresponding to it sends these families to covering families. Thus we obtain

Proposition 3.1.10 *With the above notation, the classifying topos $\mathbf{Set}[\mathbb{T}]$ may be identified with $\mathbf{Sh}(\mathcal{C}, J_{\mathbb{T}})$ where $\mathcal{C}$ is the syntactic category of $\mathbb{T}_0$ (equivalently,*

the opposite of the category of finitely presented $\mathbb{T}_0$*-models in* **Set**), *and* $J_{\mathbb{T}}$ *is the smallest coverage on this category for which the sieves generated by the families described above are* $J_{\mathbb{T}}$*-covering.* □

As with 3.1.7, this result is most useful when the coverage $J_{\mathbb{T}}$ can be explicitly determined and shown to be subcanonical. We give a couple of examples:

Examples 3.1.11 (*a*) If $\mathbb{T}$ is the theory of local rings, as described in 1.1.7(*g*), then it is natural to take $\mathbb{T}_0$ to be the (algebraic) theory of rings (throughout these examples, the word 'ring' will mean 'commutative ring with 1'). In this case, the resulting coverage on the opposite of the category $\mathbf{Rng}_\omega$ of finitely presented rings has an explicit description: it is exactly the *Zariski coverage* Z described in A2.1.11(*f*), i.e. the covering sieves are those which contain finite families of ring homomorphisms of the form $(A \to A[s_i^{-1}] \mid 1 \leq i \leq n)$, where $\{s_1, \ldots, s_n\}$ is a set of elements of A contained in no proper ideal of A (equivalently, such that there exist elements $a_1, \ldots, a_n$ of A with $\sum_{i=1}^n a_i s_i = 1$). It is also a well-known result that this coverage is subcanonical; so the generic local ring is simply the forgetful functor $\mathbf{Rng}_\omega \to \mathbf{Set}$, regarded as a ring object in the topos $\mathbf{Sh}(\mathbf{Rng}_\omega^{\mathrm{op}}, Z)$. In particular, we note that it inherits the functional completeness property of 3.1.3 from the generic ring. (We shall see some applications of this fact in Section F1.1.)

(*b*) Similarly, if $\mathbb{T}$ is the (coherent) theory of fields described in 1.1.7(*h*), we might take $\mathbb{T}_0$ to be the theory of rings. If we did so, then our coverage on $\mathbf{Rng}_\omega$ would have the property that, for any finitely presented ring A and any $s \in A$, the two morphisms $A \to A[s^{-1}]$ and $A \to A/(s)$ (where (s) is the principal ideal generated by s) would generate a covering sieve. Since both these morphisms are epic in $\mathbf{Rng}_\omega$ and their pushout is the degenerate ring 1 (which is covered by the empty sieve, because of the nontriviality axiom), it follows that if a functor $F\colon \mathbf{Rng}_\omega \to \mathbf{Set}$ is a sheaf then we have $F(A) \cong F(A[s^{-1}]) \times F(A/(s))$ for all such A and s. But the representable functors do not in general satisfy this condition (for the forgetful functor, it holds iff s is idempotent), and so the coverage is not subcanonical.

Nevertheless, there is a cartesian theory $\mathbb{T}_0$ such that the classifying topos for fields may be presented using a subcanonical coverage on the syntactic category of $\mathbb{T}_0$, namely the theory of (von Neumann) regular rings. This is obtained from the theory of rings by adding the axiom

$$(\top \vdash_x (\exists y)((x^2 y = x) \wedge (x y^2 = y)))\ .$$

It is an easy exercise in ring theory to verify that the latter axiom is cartesian relative to the theory of reduced rings (rings without nonzero nilpotent elements), and that it implies the Horn sequent $((x^2 = 0) \vdash_x (x = 0))$ which says that a ring is reduced. For more details, see [520, V 2.6].

For geometric theories $\mathbb{T}$, we may proceed in the same way as for coherent ones: that is, we may equip the geometric category $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ with its canonical coverage (consisting of all sieves generated by small covering families, cf. C2.1.12(*e*))

and then show as in 3.1.9 that the topos of sheaves on this site is a classifying topos for $\mathbb{T}$. Alternatively, we could proceed as in 3.1.10, by finding an appropriate cartesian theory $\mathbb{T}_0$ of which $\mathbb{T}$ is a quotient, and imposing a suitable coverage on the syntactic category of $\mathbb{T}_0$. However, in this case we have a possible short cut. The ∞-pretopos $\mathcal{G}_{\mathbb{T}}$ generated by $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ (cf. 1.4.12(iii)), though not small, has a separating set of objects, namely (the images of) those objects $\{\vec{x}\,.\,\phi\}$ of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ for which ϕ is in the canonical form of 1.3.8(ii). (It is clear that these form a separating set for $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$, since they meet every isomorphism class of objects of this category; and since every object of $\mathcal{G}_{\mathbb{T}}$ is a quotient of a coproduct of objects of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$, the result easily extends to the larger category.) Hence by Giraud's Theorem C2.2.8 $\mathcal{G}_{\mathbb{T}}$ is itself a Grothendieck topos. Moreover, by C2.2.10 every geometric functor from $\mathcal{G}_{\mathbb{T}}$ to a Grothendieck topos is an inverse image functor (i.e. has a right adjoint), and so we may conclude

Proposition 3.1.12 *For any geometric theory $\mathbb{T}$, the ∞-pretopos $\mathcal{G}_{\mathbb{T}}$ of* 1.4.12(iii) *(is a topos and) has the universal property of a classifying topos* **Set**$[\mathbb{T}]$ *for $\mathbb{T}$.* □

Of course, we could view this result the other way round, and regard the passage from $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ to the topos $\mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\text{geom}})$ as an alternative construction of the ∞-pretopos generated by this geometric category. Note also that, in this case, the fact that the classifying topos determines the theory up to Morita equivalence is a triviality, since Morita equivalence of geometric theories was defined as equivalence of the ∞-pretoposes generated by their syntactic categories.

Remark 3.1.13 Let $(\mathcal{C}, J)$ be a small site whose underlying category is cartesian. In 2.1.4(g) we saw how to construct a geometric sketch $\mathbb{S}$ (and hence, via 2.2.7, a geometric theory $\mathbb{T}$) whose models in any Grothendieck topos $\mathcal{E}$ correspond to cartesian cover-preserving functors $\mathcal{C} \to \mathcal{E}$. More explicitly, $\mathbb{T}$ may be obtained from the theory of cartesian functors defined on $\mathcal{C}$, as described in 1.4.8, by adding the sequent

$$\left(\top \vdash_x \bigvee_{i\in I} (\exists y_i \colon B_i)(f_i(y_i) = x)\right)$$

for each J-covering family $(f_i \colon B_i \to A \mid i \in I)$. Combining this with C2.3.9, we see that $\mathbf{Sh}(\mathcal{C}, T)$ has the universal property of a classifying topos for $\mathbb{T}$. But since every Grothendieck topos may be generated by such a site (C2.2.8(iii)), this says that *every Grothendieck topos is (equivalent to) the classifying topos of some geometric theory.* In combination with the remarks above, we may interpret this result as saying that a Grothendieck topos is 'the same thing' as a Morita equivalence class of geometric theories – or perhaps more accurately, that the notion of Grothendieck topos represents the 'extensional essence' of the intuitive (intensional) notion of a geometric theory 'up to Morita equivalence'.

Remark 3.1.14 Suppose $\mathbb{T}$ is a geometric propositional theory as defined in 1.1.7(m). In 1.4.14 we saw that the syntactic category $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ is equivalent to

a frame (L, say), and that $\mathbb{T}$ is Morita-equivalent to the theory of completely prime filters of L. It follows that $\mathbf{Set}[\mathbb{T}] \simeq \mathcal{G}_{\mathbb{T}} \simeq \mathbf{Eff}(\mathbf{Pos}_{\infty}(L))$ may be identified with the topos $\mathbf{Set}(L) \simeq \mathbf{Sh}(X)$, where X is the locale defined by $\mathcal{O}(X) = L$ (cf. the remarks at the end of Section A3.3, and C1.3.11). Thus we may extend the equivalence of 1.2.15(m) from localic toposes to arbitrary $\mathbf{Set}$-toposes: for any such topos $\mathcal{E}$, we have

$$\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E}) \simeq \mathfrak{Top}\,(\mathcal{E}, \mathbf{Sh}(X))\,.$$

In particular, we recover the result of B4.2.12 that the classifying topos of any propositional geometric theory is localic over $\mathbf{Set}$ (and that any localic $\mathbf{Set}$-topos occurs as the classifying topos of such a theory).

Clearly, if $\mathbb{T}$ is Morita-equivalent to a propositional geometric theory (equivalently, if $\mathbf{Set}[\mathbb{T}]$ is localic over $\mathbf{Set}$), then $\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$ is an essentially small preorder for any $\mathbf{Set}$-topos $\mathcal{E}$. However, the converse is false: the following counterexample is due to M. Makkai.

Example 3.1.15 Let $\mathbb{T}$ be the theory over the signature Σ with one sort A and two infinite sequences R_n, S_n of unary relation symbols, whose axioms are $(\top \vdash_x (R_n(x) \vee S_n(x)))$ and $((R_n(x) \wedge S_n(x)) \vdash_x \bot)$ for all n, plus

$$\left(\top \vdash_{x,y} \left((x = y) \vee \bigvee_{n \geq 0} ((R_n(x) \wedge S_n(y)) \vee (S_n(x) \wedge R_n(y)))\right)\right).$$

Thus a $\mathbb{T}$-model M in a $\mathbf{Set}$-topos $\mathcal{E}$ consists of an object MA together with a complemented subobject $[\![R]\!] \rightarrowtail MA \times N$ (where N is the natural number object of $\mathcal{E}$, i.e. the countable copower of 1), satisfying a condition which in particular implies that the transpose $MA \to \Omega^N$ of the classifying map of this subobject is monic. (If $\mathcal{E}$ is Boolean, the last axiom says precisely this; but in a non-Boolean topos it is stronger.) Moreover, since a morphism $f\colon M \to M'$ of $\mathbb{T}$-models must preserve the interpretations of both the R_n and the S_n, it is easy to see that it must make the diagram

$$\begin{tikzcd} MA \arrow[r, "f_A"] \arrow[dr] & M'A \arrow[d] \\ & \Omega^N \end{tikzcd}$$

commute; thus $\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$ is equivalent to a sub-poset of $\mathrm{Sub}_{\mathcal{E}}(\Omega^N)$ (and to the whole of this poset if $\mathcal{E}$ is Boolean).

However, $\mathbf{Set}[\mathbb{T}]$ is not localic over $\mathbf{Set}$. To see this, note first that $\mathbb{T}$ has 2^c non-isomorphic models in $\mathbf{Set}$, where $c = 2^{\aleph_0}$ is the cardinality of the continuum. But 1.3.8 imples that the Lindenbaum algebra of $\mathbb{T}$ is countably generated as a

complete join-semilattice, since there are only countably many regular formulae over Σ; in other words, the corresponding locale X has a countable basis in the sense of C2.2.4(b). And a locale with a basis of cardinality κ can have at most 2^κ points, since a completely prime filter in $\mathcal{O}(X)$ is determined by the basis elements which it contains (cf. C1.2.2); so $\mathbf{Sh}(X)$ cannot be equivalent to $\mathbf{Set}[\mathbb{T}]$.

We next fulfil a promise made in 1.5.12:

Proposition 3.1.16 ('Classical completeness' for geometric theories) *Let $\mathbb{T}$ be a geometric theory, and σ a geometric sequent over the signature of $\mathbb{T}$. If σ is satisfied in all $\mathbb{T}$-models in Boolean toposes, then it is provable in $\mathbb{T}$.*

Proof By A4.5.23, there exists a surjective geometric morphism $f \colon \mathcal{B} \to \mathbf{Set}[\mathbb{T}]$, where $\mathcal{B}$ is Boolean. So σ is satisfied in the $\mathbb{T}$-model $f^*(G_{\mathbb{T}})$, and hence by 1.2.13 in $G_{\mathbb{T}}$ itself; but the latter is equivalent to provability in $\mathbb{T}$. □

In particular, it follows that if a geometric sequent σ is derivable from the axioms of a geometric theory $\mathbb{T}$ using 'classical geometric logic' (i.e. the rules of geometric logic plus the Law of Excluded Middle), then there is also a constructive derivation of σ, not using the Law of Excluded Middle.

We remark that, since every Grothendieck topos is a surjective image of a localic one (cf. C5.2.1), and since the surjection constructed in A4.5.23 is localic (cf. A4.6.2(e)), there is a possible further strengthening of 3.1.16: in fact provability in a geometric theory $\mathbb{T}$ may be determined by considering satisfaction in $\mathbb{T}$-models in *localic* Boolean toposes. This has the advantage that all such toposes satisfy the Axiom of Choice (cf. 4.5.15 below); so we obtain a further conservativity result like that just mentioned, asserting that uses of the Axiom of Choice may be eliminated from any derivation of a geometric sequent from geometric hypotheses.

To conclude this section, we digress for a while from its main theme to consider full first-order theories. Since inverse image functors are not in general Heyting functors, we cannot hope to have a classifying topos (in the sense we have considered so far) for such a theory. Nevertheless, it is of interest to ask whether there is a 'completeness theorem' for such theories in terms of a single model in a Grothendieck topos, similar to that which we derived for geometric theories from the existence of a classifying topos. The answer is yes; indeed the topos in question may be taken to have a particularly simple form, as we shall see. To show this, we need a result of which the first part, at least, could have been included in Section A1.4 (and the second in Section A2.1):

Lemma 3.1.17 *Let $\mathcal{C}$ be a small Heyting category. Then*

(i) *The Yoneda embedding $Y \colon \mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ preserves universal quantification.*

(ii) *If J denotes the coherent coverage on $\mathcal{C}$, then the Yoneda embedding $Y \colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C}, J)$ is a Heyting functor.*

Proof (i) Let $f: A \to B$ be a morphism of $\mathcal{C}$, let $A' \rightarrowtail A$ be a subobject of A and let $B' = \forall_f(A')$. We must show that if $R \rightarrowtail Y(A)$ is any subobject whose pullback along $Y(f)$ is contained in $Y(A')$, then $R \leq Y(B')$. We recall that subobjects of $Y(B)$ in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ correspond to sieves on B in $\mathcal{C}$, and that the pullback of R along $Y(f)$ is the sieve of all morphisms $g: C \to A$ such that $fg \in R$. Now if $h: D \to B \in R$, form the pullback

$$\begin{tikzcd} C \arrow[r] \arrow[d, "g"] & D \arrow[d, "h"] \\ A \arrow[r, "f"] & B \end{tikzcd} ;$$

then $g \in f^*(R)$, so by assumption it factors through $A' \rightarrowtail A$. Let $D \twoheadrightarrow I \rightarrowtail B$ be the image factorization of h; since image factorizations in $\mathcal{C}$ are stable under pullback, we have $f^*(I) \leq A'$, and hence $I \leq \forall_f(A') = B'$, as required.

(ii) It is easy to see that the argument of (i) continues to work if we replace $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ by $\mathbf{Sh}(\mathcal{C}, J)$ for any subcanonical coverage J on $\mathcal{C}$. But in A2.1.11(b) we observed that the coherent coverage on a small coherent category is always subcanonical, and also that for this particular coverage the Yoneda embedding becomes a coherent functor. Since it also preserves universal quantification, it is a Heyting functor. □

Corollary 3.1.18 *For any first-order theory $\mathbb{T}$, there is a Grothendieck topos $\mathcal{E}$ and a $\mathbb{T}$-model $N_{\mathbb{T}}$ in $\mathcal{E}$ such that the first-order sequents satisfied in $N_{\mathbb{T}}$ are exactly those provable in $\mathbb{T}$.*

Proof Take $\mathcal{E} = \mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}}, J)$ where J is the coherent coverage, and let $N_{\mathbb{T}}$ be the image of $M_{\mathbb{T}}$ under the Yoneda embedding. Since the Yoneda embedding is full and faithful, it is conservative; so by 1.2.13 and 3.1.17 it preserves and reflects the truth of arbitrary first-order sequents. Hence the result follows from 1.4.11. □

It is possible to sharpen the result of 3.1.18. We again recall from C5.2.1 that, for any Grothendieck topos $\mathcal{E}$, there exists an open surjection $f: \mathcal{F} \to \mathcal{E}$ where $\mathcal{F}$ is localic (over **Set**); and from C3.1.7 that the inverse image of an open surjection of toposes is a conservative Heyting functor. Thus we may require the topos $\mathcal{E}$ of 3.1.18 to be localic over **Set**. In fact P. Freyd [381] has shown, by a substantially different method, that we may require it to be the topos of sheaves on a Stone space (that is, a compact zero-dimensional space – equivalently, the prime ideal space of a Boolean algebra); and that if our signature Σ is countable (so that the category $\mathcal{C}_{\mathbb{T}}^{\mathrm{fo}}$ is also countable) we may actually take $\mathcal{E}$ to be the particular topos $\mathbf{Sh}(K)$, where K is the Cantor space. (In other words, $\mathbf{Sh}(K)$ plays a rôle in constructive first-order logic analogous to that of **Set** in classical first-order logic.)

It is also possible to generalize 3.1.18 to infinitary first-order logic, as defined in 1.1.3(g). However, in contrast to the coherent/geometric case, we cannot work with the whole of infinitary first-order logic in constructing our syntactic category: the reason is that the resulting category will not in general be small, or even equivalent to a small category. (Even in the case of a propositional signature with no axioms, the syntactic category would be equivalent to the free complete Heyting algebra on the set of atomic propositions, and if there is more than one atomic proposition then the latter is known to be a proper class; cf. [550].) We must therefore restrict ourselves to *κ-first-order logic* for some infinite regular cardinal κ, that is the fragment in which we allow conjunctions and disjunctions indexed by sets of cardinality less than κ, but not any larger ones. (Provided we retain the convention that a theory is specified by a *set* of axioms, it is clear that for any infinitary first-order theory $\mathbb{T}$ we can find a cardinal κ such that $\mathbb{T}$ lies in κ-first-order logic.) Given a theory $\mathbb{T}$ in κ-first-order logic, we may build its syntactic category $\mathcal{C}_{\mathbb{T}}^{\kappa\text{fo}}$ in the usual way, and verify that it is a (small) κ-Heyting category (that is, a Heyting category having unions and intersections of families of subobjects of cardinality less than κ, which are stable under pullback), and that κ-Heyting functors from $\mathcal{C}_{\mathbb{T}}^{\kappa\text{fo}}$ to another such category $\mathcal{D}$ correspond to $\mathbb{T}$-models in $\mathcal{D}$. If we equip $\mathcal{C}_{\mathbb{T}}^{\kappa\text{fo}}$ with the coverage J_κ whose covering families are '$\mathbb{T}$-provably surjective' families of cardinality less than κ, then we may further verify that J_κ is subcanonical and that the Yoneda embedding $\mathcal{C}_{\mathbb{T}}^{\kappa\text{fo}} \to \mathbf{Sh}(\mathcal{C}_{\mathbb{T}}^{\kappa\text{fo}}, J_\kappa)$ is a κ-Heyting functor. Thus we obtain

Proposition 3.1.19 *Let $\mathbb{T}$ be any theory in κ-first-order logic. Then there exists a Grothendieck topos $\mathcal{E}$ (which may be taken to be localic over* **Set***) containing a $\mathbb{T}$-model which satisfies exactly the κ-first-order sequents which are provable in $\mathbb{T}$.* □

However, this construction is not universal in any sense: in particular, different choices of the cardinal κ will yield inequivalent toposes. For more details, see [212].

Suggestions for further reading: Butz & Johnstone [212], Cole [245], Dubuc & Reyes [309], Tierney [1169], Wraith [1246].

D3.2 The object classifier

In this section we shall study one particular classifying topos: namely that for the theory of *objects*, i.e. the empty theory over the signature with one sort and no primitive symbols except equality. (As in Section B4.2, we shall denote this theory by $\mathbb{O}$.) Clearly, there is not very much of interest to be said about $\mathbb{O}$ as a theory; but its classifying topos $\mathbf{Set}[\mathbb{O}]$ has some interesting structure, which it seems appropriate to explore at this point.

By 3.1.2, $\mathbf{Set}[\mathbb{O}]$ may be identified with the functor category $[\mathbf{Set}_f, \mathbf{Set}]$, where $\mathbf{Set}_f$ is the category of finite sets (or a small skeleton thereof, such as the

category of finite cardinals and all maps between them), and the generic object $G_{\mathbb{O}}$ itself is just the inclusion functor $\mathbf{Set}_f \to \mathbf{Set}$.

Given an arbitrary object A of a **Set**-topos $\mathcal{E}$, we have a corresponding geometric morphism $\mathcal{E} \to [\mathbf{Set}_f, \mathbf{Set}]$, which we shall denote by $\overline{A}$, and which is characterized by the property that $\overline{A}^*(G_{\mathbb{O}}) \cong A$. Explicitly, A corresponds to the $\mathbf{Set}_f$-torsor $h_A \colon \mathbf{Set}_f^{\mathrm{op}} \to \mathcal{E}$ given by $h_A(n) = A^n$, and $\overline{A}^*$ sends a functor $T \colon \mathbf{Set}_f \to \mathbf{Set}$ to the tensor product $h_A \otimes_{\mathbf{Set}_f} T$, i.e. to the coequalizer of

$$\coprod_{\alpha : m \to n, x \in T(m)} A^n \rightrightarrows \coprod_{n \in \mathbb{N}, x \in T(n)} A^n$$

where the maps are induced respectively by the actions of $\mathbf{Set}_f$ on T and on h_A. Equivalently, we can describe $\overline{A}^*(T)$ as the coend $\int^n T(n) \times A^n$.

In particular, given an object T of $\mathbf{Set}[\mathbb{O}]$, we have a geometric morphism $\overline{T} \colon \mathbf{Set}[\mathbb{O}] \to \mathbf{Set}[\mathbb{O}]$ (unique up to isomorphism) such that $\overline{T}^*(G_{\mathbb{O}}) \cong T$. Composition of geometric endomorphisms of $\mathbf{Set}[\mathbb{O}]$ gives rise to a bifunctor $\otimes \colon \mathbf{Set}[\mathbb{O}] \times \mathbf{Set}[\mathbb{O}] \to \mathbf{Set}[\mathbb{O}]$, defined by

$$S \otimes T = \overline{T}^* \overline{S}^* (G_{\mathbb{O}}) ,$$

so that $\overline{S \otimes T} \cong \overline{S} \cdot \overline{T}$. The properties of this bifunctor are summarized by

Proposition 3.2.1 $(\mathbf{Set}[\mathbb{O}], \otimes, G_{\mathbb{O}})$ *is a (non-symmetric) monoidal category, which is closed on the right, i.e. for each object T the functor $(-) \otimes T$ has a right adjoint.*

Proof The associativity and unit isomorphisms for $\otimes$ are obtained by transposing the associativity and unit identities for composition of geometric morphisms across the equivalence

$$\mathfrak{Top}/\mathbf{Set}\,(\mathbf{Set}[\mathbb{O}], \mathbf{Set}[\mathbb{O}]) \simeq \mathbf{Set}[\mathbb{O}] .$$

It is straightforward to verify that they satisfy the coherence conditions for a monoidal structure. The second assertion follows from the fact that $(-) \otimes T$ is isomorphic to the functor $\overline{T}^*$, which has a right adjoint $\overline{T}_*$. □

What can we say about the monoids for this monoidal structure? Clearly, an object T of $\mathbf{Set}[\mathbb{O}]$ induces an endofunctor $T_{\mathcal{E}}$ of every **Set**-topos $\mathcal{E}$, defined by $T_{\mathcal{E}}(A) = \overline{A}^*(T)$ $(\cong \overline{A}^* \overline{T}^* (G_{\mathbb{O}}))$. If T has a monoid structure, then $T_{\mathcal{E}}$ has a monad structure, which is 'natural' with respect to arbitrary inverse image functors between **Set**-toposes. We shall see in Section D5.3 below that such monads may be naturally identified with finitary algebraic theories.

We may give an explicit description of the tensor product, as follows. Specializing the coend formula above for $\overline{A}^*(T)$ to the case when $\mathcal{E} = \mathbf{Set}[\mathbb{O}]$ and A

is a functor $S\colon \mathbf{Set}_f \to \mathbf{Set}$, we obtain

$$(T \otimes S)(m) = \int^n T(n) \times (S(m))^n \ ;$$

that is, $(T\otimes S)(m)$ is the quotient of the disjoint union of the sets $T(n)\times(S(m))^n$ by the smallest equivalence relation which identifies $(Tf(x), (y_1, \ldots, y_n))$ with $(x, (y_{f1}, \ldots, y_{fn'}))$ whenever $f\colon n' \to n$ in $\mathbf{Set}_f$, $x \in T(n')$ and $y_1, \ldots, y_n \in S(m)$. It is easy to see directly that, if T is (the restriction to $\mathbf{Set}_f$ of) the functor underlying a finitary monad on **Set** (i.e. a finitary algebraic theory), then it carries a $\otimes$-monoid structure: for we may identify the elements of $T(n)$ with n-ary operations of the theory $\mathbb{T}$, and the multiplication sends $(\alpha, (\beta_1, \ldots, \beta_n))$ to the composite operation

$$(z_1, \ldots, z_m) \mapsto \alpha(\beta_1(z_1, \ldots, z_m), \ldots, \beta_n(z_1, \ldots, z_m)) \ .$$

(Similarly, the unit map $G_{\mathbb{O}} \to T$ sends the ith element of $n = G_{\mathbb{O}}(n)$ to the ith projection, considered as an n-ary operation of $\mathbb{T}$.)

We note that the generic object $G_{\mathbb{O}}$ is a functor taking finite sets as values. This does not imply that, as an object of $\mathbf{Set}[\mathbb{O}]$, it is finite in the Kuratowskian sense defined in 1.1.7(k); for K-finiteness is preserved by inverse image functors, and so if $G_{\mathbb{O}}$ were K-finite then every object of every **Set**-topos would be K-finite. However, it does satisfy another, non-geometric, notion of finiteness, namely the sequent

$$(\neg(x = y) \vdash_{x,y} \bot),$$

which says that it does not have two disjoint elements. For the interpretation of this sequent is that the pseudo-complement of the diagonal subobject of $G_{\mathbb{O}} \times G_{\mathbb{O}}$ is empty, and this is easily seen to be true since any nonempty subobject of $G_{\mathbb{O}} \times G_{\mathbb{O}}$ must contain the unique element of $(G_{\mathbb{O}} \times G_{\mathbb{O}})(1)$. On the other hand, $G_{\mathbb{O}}$ is also 'potentially infinite' in the sense that, for each natural number m, it satisfies the non-geometric sequent

$$((\forall x)((x = y_1) \vee \cdots \vee (x = y_m)) \vdash_{y_1, \ldots, y_m} \bot)$$

which says that it cannot be 'covered by m elements'. (We interpret the case $m = 0$ of this sequent as $((\forall x)\bot \vdash_{[]} \bot)$, i.e. the assertion that $G_{\mathbb{O}}$ is nonempty in the 'negative' sense, as opposed to the positive sense $(\top \vdash_{[]} (\exists x)\top)$.) More generally, it may be shown that $G_{\mathbb{O}}$ satisfies

$$((\forall x)((x = y_1) \vee \cdots \vee (x = y_m) \vee \phi) \vdash_{y_1,\ldots,y_m} \phi)$$

where ϕ is any coherent formula with $\mathrm{FV}(\phi) \subseteq \{y_1, \ldots, y_m\}$. (The previous displayed sequent is of course the special case $\phi = \bot$ of this scheme.) We may therefore conclude that the above sequents are 'geometrically inconsequential',

in the following sense:

Lemma 3.2.2 *Let σ be a geometric sequent over the signature Σ with one sort and no primitive symbols except equality, and suppose σ is (constructively) provable in the theory $\mathbb{T}$ whose axioms are the sequent $(\neg(x = y) \vdash_{x,y} \bot)$ and all instances of the last displayed sequent on the previous page. Then σ is provable in the empty theory over Σ.* □

In fact it can be shown (cf. [212]) that the theory $\mathbb{T}$ of 3.2.2 is the full (infinitary) first-order theory of $G_{\mathbb{O}}$: that is, any sequent satisfied by $G_{\mathbb{O}}$ is deducible from $\mathbb{T}$. (Equivalently, if A is any $\mathbb{T}$-model in a Grothendieck topos $\mathcal{E}$, then $\overline{A}\colon \mathcal{E} \to \mathbf{Set}[\mathbb{O}]$ is an open map, so that A satisfies all the infinitary first-order sequents satisfied by $G_{\mathbb{O}}$.)

The study of such 'inconsequential' non-geometric properties of generic models is not entirely a pointless curiosity. Although 3.2.2 has no known applications, we shall see in Section F1.1 how the fact that the generic local ring satisfies a suitable non-geometric 'field axiom' enables us to establish results about projective geometry over the generic local ring, and then (since these results are expressible by geometric sequents) to 'transport' them to results about projective geometry over an arbitrary local ring.

An interesting fact about $\mathbf{Set}[\mathbb{O}]$ is that every Grothendieck topos $\mathcal{E}$ admits a localic geometric morphism to it. In order to prove this, we shall introduce a weakening of the notion of a bound for a geometric morphism, which we met in Section B3.1.

We recall from B3.1.7 that an object B of an $\mathcal{S}$-topos $p\colon \mathcal{E} \to \mathcal{S}$ is called a *bound* (for $\mathcal{E}$ over $\mathcal{S}$) if every object of $\mathcal{E}$ is a subquotient of one of the form $p^*I \times B$, where I is an object of $\mathcal{S}$. When $\mathcal{S} = \mathbf{Set}$, by B3.1.8(*b*), this is equivalent to saying that the subobjects of B form a separating set for $\mathcal{E}$. The generic object $G_{\mathbb{O}}$ is not a bound for $\mathbf{Set}[\mathbb{O}]$ over $\mathbf{Set}$; but it is 'very nearly so', in the following sense. By the Yoneda lemma, the representable functors $\mathbf{Set}_f(n, -)$ form a separating set for $\mathbf{Set}[\mathbb{O}]$, so their coproduct forms a bound for it; but these functors are just the finite powers of the generic object $\mathbf{Set}_f(1, -)$, so their coproduct is the list object $L(G_{\mathbb{O}})$ over $G_{\mathbb{O}}$, as defined in A2.5.17.

Definition 3.2.3 Given a geometric morphism $p\colon \mathcal{E} \to \mathcal{S}$, where $\mathcal{S}$ (and hence $\mathcal{E}$) has a natural number object, we shall say that an object B of $\mathcal{E}$ is a *pre-bound* for $\mathcal{E}$ over $\mathcal{S}$ if the list object LB is a bound for $\mathcal{E}$ over $\mathcal{S}$.

If $\mathcal{S}$ is the topos $\mathbf{Set}$, this is equivalent to saying that the subobjects of finite powers of B form a separating set for $\mathcal{E}$: one direction is trivial, and the converse follows from the fact that any subobject of LB decomposes as a coproduct of subobjects of finite powers of B. We recall that any object which contains a bound as a subobject (or has one as an epimorphic image) is itself a bound; hence in particular any bound is a pre-bound, since LB contains B as a subobject.

Lemma 3.2.4 *For a geometric morphism $f: \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f is localic.*

(ii) *For any geometric morphism $p: \mathcal{E} \to \mathcal{S}$ and any bound B for $\mathcal{E}$ over $\mathcal{S}$, f^*B is a bound for $\mathcal{F}$ over $\mathcal{S}$.*

(iii) *There exists a geometric morphism $p: \mathcal{E} \to \mathcal{S}$ and an object B of $\mathcal{E}$ such that f^*B is a bound for $\mathcal{F}$ over $\mathcal{S}$.*

(iv) *(if $\mathcal{E}$ has a natural number object) For any geometric morphism $p: \mathcal{E} \to \mathcal{S}$ where $\mathcal{S}$ has a natural number object, and any pre-bound B for $\mathcal{E}$ over $\mathcal{S}$, f^*B is a pre-bound for $\mathcal{F}$ over $\mathcal{S}$.*

(v) *(if $\mathcal{E}$ has a natural number object) There exists a geometric morphism $p: \mathcal{E} \to \mathcal{S}$ where $\mathcal{S}$ has a natural number object, and an object B of $\mathcal{E}$, such that f^*B is a pre-bound for $\mathcal{F}$ over $\mathcal{S}$.*

Proof (i) $\Rightarrow$ (ii) follows easily from inspection of the proof, given in B3.1.10(i), that a composite of bounded geometric morphisms is bounded.

(ii) $\Rightarrow$ (iii) is trivial (take p to be the identity morphism on $\mathcal{E}$, and B to be its terminal object).

(iii) $\Rightarrow$ (i): by assumption, every object of $\mathcal{F}$ is a subquotient of one of the form $(pf)^*I \times f^*B$, for some $I \in \text{ob}\, \mathcal{S}$. But this is isomorphic to $f^*(p^*I \times B)$, and hence in the image of f^*; so f is localic.

(ii) $\Rightarrow$ (iv) follows easily from the fact that f^* commutes up to isomorphism with the list-object functor, as we remarked after A2.5.15.

(iv) $\Rightarrow$ (v) is similar to (ii) $\Rightarrow$ (iii) (and (iii) $\Rightarrow$ (v) is also trivial). And (v) $\Rightarrow$ (i) is similar to (iii) $\Rightarrow$ (i), again using the fact that f^* commutes with the list-object functor. □

Theorem 3.2.5 *Let $\mathcal{E}$ be a Grothendieck topos, B an object of $\mathcal{E}$. The following are equivalent:*

(i) *B is a pre-bound for $\mathcal{E}$ over* **Set**.

(ii) *The geometric morphism $\overline{B}: \mathcal{E} \to$* **Set**$[\mathbb{O}]$ *corresponding to B is localic.*

(iii) *There exists a single-sorted geometric theory $\mathbb{T}$ and an equivalence $\mathcal{E} \simeq$* **Set**$[\mathbb{T}]$ *which identifies B with the underlying object of the generic $\mathbb{T}$-model $G_{\mathbb{T}}$.*

Proof (i) $\Leftrightarrow$ (ii) follows immediately from 3.2.4 and the fact, noted above, that $G_{\mathbb{O}}$ is a pre-bound for **Set**$[\mathbb{O}]$ over **Set**.

(iii) $\Rightarrow$ (ii): Suppose (iii) holds. Then, given an arbitrary Grothendieck topos $\mathcal{F}$ and an object A of $\mathcal{F}$, factorizations of the classifying map $\overline{A}: \mathcal{F} \to$ **Set**$[\mathbb{O}]$ through $\overline{B}$ correspond to $\mathbb{T}$-model structures on the given object A of $\mathcal{F}$. In other words, as a topos over **Set**$[\mathbb{O}]$, $\mathcal{E} \simeq$ **Set**$[\mathbb{T}]$ becomes the classifying topos for a propositional geometric theory in the sense of 1.1.7(m) – the theory of $\mathbb{T}$-model structures on the *particular* object $G_{\mathbb{O}}$ of **Set**$[\mathbb{O}]$. So by B4.2.12 the geometric morphism $\overline{B}$ is localic.

Alternatively, we could give a direct proof of (iii) $\Rightarrow$ (i), by observing that in the syntactic category $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ every object is a subobject of a finite power of (the underlying object of) $M_{\mathbb{T}}$; but by 3.1.12 $\mathbf{Set}[\mathbb{T}]$ may be taken to be the ∞-pretopos generated by $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$, and so every object of it is a quotient of a coproduct of objects of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$. So the subobjects of finite powers of the underlying object of $G_{\mathbb{T}}$ form a separating set for it.

(i) $\Rightarrow$ (iii): We take a single-sorted signature Σ_B having an n-ary relation symbol $[R]$ for each subobject $R \rightarrowtail B^n$ in $\mathcal{E}$. (We could also add n-ary function symbols corresponding to the morphisms $B^n \to B$ in $\mathcal{E}$, but since their graphs will appear as $(n+1)$-ary relation symbols, this is not necessary.) Then B itself carries an obvious Σ_B-structure in $\mathcal{E}$; we define $\mathbb{T}$ to be the set of all geometric sequents over Σ_B which are satisfied in this structure. Thus B is a $\mathbb{T}$-model; we shall show that the geometric morphism $\mathcal{E} \to \mathbf{Set}[\mathbb{T}]$ which corresponds to it is an equivalence.

Consider first the geometric functor $F_B : \mathcal{C}_{\mathbb{T}}^{\text{geom}} \to \mathcal{E}$ which corresponds to B. Since every object $\{x_1, \ldots, x_n . \phi\}$ of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ is a subobject of $\{x_1, \ldots, x_n . \top\} \cong (M_{\mathbb{T}})^n$ for some n, every object in the image of F_B is a subobject of a finite power of B; but the converse is also true, by the construction of Σ_B. Also, if $f : R \to S$ is any morphism of $\mathcal{E}$ between subobjects of finite powers of B, then the sequents which say that (the formula over Σ_B corresponding to) the graph of f represents a morphism of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ are satisfied in B and hence provable in $\mathbb{T}$; so F_B is full. Similarly, if two parallel morphisms $[\theta]$ and $[\theta']$ of $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ have the same image under F_B, then θ and θ' must be provably equivalent in $\mathbb{T}$; so F_B is faithful. Since F_B is a geometric functor, it carries covering families in $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ to epimorphic families in $\mathcal{E}$; but, once again, the definition of $\mathbb{T}$ ensures that the converse is true. Finally, since B is a pre-bound for $\mathcal{E}$ over $\mathbf{Set}$, the objects in the image of F_B form a separating set for $\mathcal{E}$; so we are in a position to apply the Comparison Lemma (C2.2.3) and conclude that $\mathcal{E}$ is equivalent to the topos of sheaves on $\mathcal{C}_{\mathbb{T}}^{\text{geom}}$ for the canonical coverage. But the latter is exactly the classifying topos $\mathbf{Set}[\mathbb{T}]$. $\square$

Since, by definition, every Grothendieck topos has a bound (and hence a pre-bound) over $\mathbf{Set}$, we obtain as an immediate corollary the result, promised above, that every such topos admits a localic morphism to $\mathbf{Set}[\mathbb{O}]$. (We also obtain a new proof of 1.4.13; however, 1.4.13 could alternatively have been used to prove this result, by the argument given in the proof of (iii) $\Rightarrow$ (ii) above.)

We note the following easy consequences of 3.2.5:

Corollary 3.2.6 *Let $\mathbb{T}$ be a single-sorted geometric theory, and M a $\mathbb{T}$-model in a Grothendieck topos $\mathcal{E}$, classified by a geometric morphism $f : \mathcal{E} \to \mathbf{Set}[\mathbb{T}]$. Then*

(i) *f is localic iff the underlying object of M is a pre-bound for $\mathcal{E}$ over* $\mathbf{Set}$.

(ii) *f is surjective iff M is a conservative $\mathbb{T}$-model (i.e. satisfies only those geometric sequents derivable from $\mathbb{T}$).*

Proof (i) Let $g\colon \mathbf{Set}[\mathbb{T}] \to \mathbf{Set}[\mathbb{O}]$ classify the underlying object of the generic $\mathbb{T}$-model. Then g is localic by 3.2.5, so by A4.6.2(e) and (f) the morphism f is localic iff the composite gf is. But gf classifies the underlying object of M, so the result is immediate from 3.2.5.

(ii) One direction is immediate from 1.2.13: if f is surjective, then any geometric sequent satisfied in M is satisfied in $G_{\mathbb{T}}$, and hence provable in $\mathbb{T}$. Conversely, suppose M is a conservative $\mathbb{T}$-model, and consider the class $\mathcal{C}$ of objects A of $\mathbf{Set}[\mathbb{T}]$ such that f^* maps proper subobjects of A to proper subobjects of f^*A. Conservativity of M says that $\mathcal{C}$ contains all subobjects of finite powers of $G_{\mathbb{T}}$; but since f^* preserves coproducts and epimorphisms, it is easy to see that $\mathcal{C}$ must be closed under coproducts and quotients, and hence (since $G_{\mathbb{T}}$ is a prebound for $\mathbf{Set}[\mathbb{T}]$) that it is the class of all objects of $\mathbf{Set}[\mathbb{T}]$. So by A1.2.4 f^* is conservative. □

We conclude this section by describing the classifying toposes of two theories fairly closely related to $\mathbb{O}$: the theory $\mathbb{D}$ of decidable objects and the theory $\mathbb{K}$ of K-finite objects. $\mathbb{D}$ is written over the signature Σ with one sort and one binary relation # apart from equality (we write this relation in infix rather than prefix notation, i.e. $(x \# y)$ instead of $\#(x,y)$).The axioms of $\mathbb{D}$ are the sequents

$$((x \# x) \vdash_x \bot)$$

and

$$(\top \vdash_{x,y} ((x \# y) \vee (x = y)))$$

which say that in any $\mathbb{D}$-model A the interpretation of # is the complement of the diagonal subobject of $A \times A$. Thus, for any topos $\mathcal{E}$, $\mathbb{D}$-$\mathbf{Mod}(\mathcal{E})$ is the category of decidable objects of $\mathcal{E}$ (that is, objects with complemented diagonal) and monomorphisms between them. (Recall from B4.2.3 that we cannot have a geometric theory whose category of models in an arbitrary $\mathcal{E}$ is the category of all objects of $\mathcal{E}$ and monomorphisms between them; $\mathbb{D}$ achieves this for all Boolean toposes $\mathcal{E}$, which is best possible.)

We may construct a classifying topos for $\mathbb{D}$ using the method of 3.1.10. For this purpose, we could take $\mathbb{D}_0$ to be the empty theory over Σ, but it will simplify matters slightly if we allow it to include the Horn sequent

$$((x \# y) \vdash_{x,y} (y \# x))$$

which says that # is symmetric (it is clear that this sequent is provable in $\mathbb{D}$). Thus $\mathbb{D}_0$-models are objects equipped with a symmetric binary relation: what are sometimes called *undirected graphs* (with possible 'loops' joining a vertex to itself, but without multiple edges). The classifying topos for $\mathbb{D}_0$ is clearly $[\mathbf{Gph}_f, \mathbf{Set}]$, where $\mathbf{Gph}_f$ is the category of finite undirected graphs and graph homomorphisms. The axioms of $\mathbb{D}$ correspond to a coverage on $\mathbf{Gph}_f^{\mathrm{op}}$ in which the graph with one vertex and a loop is covered by the empty family (and hence,

by pushout, *any* finite graph with a loop is covered by the empty family), and the discrete graph 2 with two vertices is covered by the pair consisting of the inclusion $2 \to \mathbf{2}$ (where $\mathbf{2}$ is the connected graph with two vertices) and the unique morphism $2 \to 1$. By pushing out and composing copies of these covering families, we may deduce that, for any finite graph $(A, \#)$, we have a covering family of morphisms

$$((A, \#) \to (A/E, \#_E) \mid E \in I) ,$$

where I is the set of all equivalence relations on A which are disjoint from $\#$ (note that this set is empty if $\#$ meets the diagonal), A/E is the quotient of A by the equivalence relation E, and $\#_E$ is the complement of the diagonal relation on A/E. But it is easy to check that, if we define a cosieve on $(A, \#)$ to be covering whenever it contains the above family, then we do obtain a Grothendieck coverage on $\mathbf{Gph}_f^{\mathrm{op}}$; so this is the coverage $J_{\mathbb{D}}$ which we require. Further, we see that if an object $(A, \#)$ has the property that $\#$ is the complement of the diagonal (i.e. if it is a model of $\mathbb{D}$), then its only covering cosieve is the maximal one (cf. C2.2.18(a)). So, if we identify the category $\mathbf{Set}_{fm}$ of finite sets and injections with the full subcategory of $\mathbf{Gph}_f$ on these objects, then by C2.2.4(d) we have an equivalence of categories

$$[\mathbf{Set}_{fm}, \mathbf{Set}] \simeq \mathbf{Sh}(\mathbf{Gph}_f^{\mathrm{op}}, J_{\mathbb{D}}) .$$

Thus we have proved

Proposition 3.2.7 *The classifying topos* $\mathbf{Set}[\mathbb{D}]$ *for the coherent theory* $\mathbb{D}$ *of decidable objects may be taken to be the functor category* $[\mathbf{Set}_{fm}, \mathbf{Set}]$. □

The generic decidable object $G_{\mathbb{D}}$ is of course the inclusion functor $\mathbf{Set}_{fm} \to \mathbf{Set}$; recall that an object F of a functor category $[\mathcal{C}, \mathbf{Set}]$ is decidable iff $F(\alpha)$ is injective for each morphism α of $\mathcal{C}$ (A1.4.16(a)).

Remark 3.2.8 The fact that the classifying topos for $\mathbb{D}$ is a functor category is not unrelated to the fact that $\mathbb{D}$ is a disjunctive theory in the sense of 1.3.6. It is not true that every disjunctive theory has a functor category as its classifying topos (this fails, for example, for the theory of fields as described in 1.1.7(h)); but for every such theory $\mathbb{T}$ over a signature Σ, there is a disjunctive theory $\mathbb{T}'$ over a larger signature Σ' (sometimes called the *Diers completion* of $\mathbb{T}$) whose classifying topos is a functor category, and such that the forgetful functor $\Sigma'\text{-}\mathbf{Str}(\mathbf{Set}) \to \Sigma\text{-}\mathbf{Str}(\mathbf{Set})$ restricts to an isomorphism $\mathbb{T}'\text{-}\mathbf{Mod}(\mathbf{Set}) \cong \mathbb{T}\text{-}\mathbf{Mod}(\mathbf{Set})$. (In the case when $\mathbb{T}$ is coherent, i.e. involves only finite disjunctions, we may replace $\mathbf{Set}$ by an arbitrary Boolean Grothendieck topos in the above.) For more details, see [508].

We mention in passing that the method of constructing a classifying topos for a theory $\mathbb{T}$ by way of a 'simpler' theory $\mathbb{T}_0$ of which it is a quotient, which

we sketched in 3.1.7 and 3.1.10, will continue to work even if $\mathbb{T}_0$ is not cartesian, provided we have a suitable representation of $\mathbf{Set}[\mathbb{T}_0]$ as a functor category $[\mathcal{C}, \mathbf{Set}]$. We shall exploit this observation, with $\mathbb{T}_0 = \mathbb{D}$, in 3.4.10 below.

Remark 3.2.9 An alternative proof of 3.2.7 could be given 'from the opposite end', by starting with a presentation of the geometric theory classified by $[\mathbf{Set}_{fm}, \mathbf{Set}]$ (that is, the theory of torsors on $\mathbf{Set}_{fm}$), and showing that it is Morita-equivalent to $\mathbb{D}$. This is done, for example, in [549].

We gave an explicit presentation of the theory $\mathbb{K}$ of K-finite objects in 1.1.7(k). Once again, we approach the problem of describing its classifying topos by the method of 3.1.10: this time we take $\mathbb{K}_0$ to be the Horn theory specified by (all instances of) the last two displayed sequents of 1.1.7(k). Thus a $\mathbb{K}_0$-model is a set A equipped with a notion of 'distinguished n-tuple' of elements for every $n > 0$, such that if $(a_1, \ldots, a_n)$ is a distinguished n-tuple then so is any m-tuple obtained from it by re-ordering and either introducing or removing duplication of elements. Equivalently, therefore, we can think of it as a set equipped with a family of (nonempty) distinguished finite subsets. (In addition, a structure for the appropriate signature involves a truth-value $[\![R_0]\!]$, about which the theory $\mathbb{K}_0$ makes no assertions; we can think of R_0 as the assertion that the empty subset is distinguished.) Once again, it is clear that the finitely presented $\mathbb{K}_0$-models are just the finite ones; let us write $\mathcal{C}$ for the category of finite $\mathbb{K}_0$-models.

The case $n = 0$ of the first displayed sequent of 1.1.7(k) tells us that the coverage $J_{\mathbb{K}}$ on $\mathcal{C}^{\mathrm{op}}$ must include the empty covering cosieve on the object given by a singleton with the empty set as the only distinguished subset; it follows that any nonempty finite $\mathbb{K}_0$-model with the empty set distinguished is covered by the empty cosieve, and so we can effectively forget about these objects (cf. C2.2.4(e)). For $n > 0$, the first sequent of 1.1.7(k) tells us that, if A is the set $\{a_1, \ldots, a_n\}$ with the whole of A as the only distinguished subset, and B is obtained from A by adding a new element b but retaining A as the only distinguished subset, then the family of all splittings for the inclusion $A \subseteq B$ generates a covering cosieve on B. It follows that, for any object C of $\mathcal{C}$ with at least one nonempty distinguished subset, the family of all surjections $f \colon C \twoheadrightarrow D$, where the only distinguished subset of D is the whole of D, generates a covering cosieve. Finally, the second displayed sequent of 1.1.7(k) says that the initial object of $\mathcal{C}$ (the empty $\mathbb{K}_0$-model with no distinguished subsets) is covered by the cosieve consisting of the identity map to the empty $\mathbb{K}_0$-model with $\emptyset$ distinguished, together with all maps to nonempty $\mathbb{K}_0$-models with at least one distinguished subset. So, as with the theory $\mathbb{D}$, we see that a cosieve on an arbitrary object A of $\mathcal{C}$ is $J_{\mathbb{K}}$-covering iff it contains all morphisms with domain A whose codomain lies in the full subcategory (isomorphic to the category $\mathbf{Set}_{fe}$ of finite sets and surjections between them) whose objects are the (finite) $\mathbb{K}$-models. Applying C2.2.4(d) once

again, we deduce

Proposition 3.2.10 *The classifying topos* $\mathbf{Set}[\mathbb{K}]$ *for the geometric theory* $\mathbb{K}$ *of* K*-finite objects may be taken to be the functor category* $[\mathbf{Set}_{fe}, \mathbf{Set}]$. □

As before, the generic $\mathbb{K}$-model is the inclusion functor $\mathbf{Set}_{fe} \to \mathbf{Set}$; it follows from 5.4.14 below that this is a K-finite object of $[\mathbf{Set}_{fe}, \mathbf{Set}]$. And, once again, 3.2.10 could alternatively have been proved by obtaining a presentation of the theory classified by $[\mathbf{Set}_{fe}, \mathbf{Set}]$ and showing that it is Morita-equivalent to $\mathbb{K}$; see [549].

Suggestions for further reading: Johnstone & Wraith [549], Ščedrov [1085].

D3.3 Coherent toposes

In this section we study the particular properties of those Grothendieck toposes which occur as the classifying toposes of coherent theories. Such toposes are called *coherent toposes*; from our point of view, this name is slightly unfortunate, since (as we saw in A2.3.5) *all* toposes are coherent categories. However, the term 'coherent topos' represents the original use of the word 'coherent' in topos theory (it was borrowed from the 'coherent sheaves' which occur in algebraic geometry), and 'coherent theory' and 'coherent category' are successive back-formations from this.

We begin with

Theorem 3.3.1 *Let* $\mathcal{E}$ *be a Grothendieck topos. The following conditions are equivalent:*

(i) *There exists a coherent (resp. regular) theory* $\mathbb{T}$ *such that* $\mathcal{E} \simeq \mathbf{Set}[\mathbb{T}]$.

(ii) *There exists a site of definition* $(\mathcal{C}, T)$ *for* $\mathcal{E}$ *such that* $\mathcal{C}$ *is small and cartesian, and* T *is generated by finite (resp. singleton) covering families.*

(iii) *There exists a standard site of definition* $(\mathcal{C}, T)$ *for* $\mathcal{E}$ *such that* T *is generated by finite (resp. singleton) covering families.*

(iv) *There exists a small pretopos (resp. a small effective regular category)* $\mathcal{C}$ *such that* $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, T)$, *where* T *is the coherent (resp. regular) coverage on* $\mathcal{C}$.

Proof (ii) $\Rightarrow$ (i) follows from the argument given in 3.1.13 to show that every Grothendieck topos is a classifying topos; for, in the axioms of the geometric theory $\mathbb{T}$ described there, the only disjunctions which appear are indexed by (generating subsets of) the covers in T. (iv) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (ii) are trivial; and (i) $\Rightarrow$ (iv) follows from the construction of $\mathbf{Set}[\mathbb{T}]$ mentioned after 3.1.9 (resp. in 3.1.5). (Note that the alternative construction of 3.1.10 (resp. 3.1.7) yields a site of definition for $\mathcal{E}$ satisfying the hypotheses of (ii), but not necessarily those of (iii).) □

Before proceeding further, we remark that the presence of finite limits in the underlying category, as well as the finiteness of the covering families, is an essential feature of the characterization in 3.3.1(ii). For example, we saw in 3.2.10 that the theory $\mathbb{K}$ of K-finite objects has a classifying topos which is a functor category (i.e. has a site of definition with no nontrivial covering families); but we also saw in 2.4.12(c) that this theory cannot be Morita-equivalent to any finitary first-order theory, let alone a coherent one.

If $\mathcal{E}$ satisfies the hypotheses of 3.3.1, we shall call it a *coherent topos* (resp. a *regular topos*). For the rest of this section, we shall deal in detail only with the coherent case, though we shall briefly discuss the corresponding results for the regular case in 3.3.10 below.

In the search for a characterization of coherent toposes, the property of compactness plays an important rôle. We say that an object A of a topos (or more generally a geometric category) is *compact* if the top element of the lattice $\mathrm{Sub}(A)$ is compact; equivalently, if every covering family $(A_i \to A \mid i \in I)$ contains a finite covering subfamily.

Example 3.3.2 In the topos $\mathbf{Sh}(X)$ of sheaves on a locale X, an object A is compact iff the domain of the corresponding local homeomorphism $E \to X$ (cf. C1.3.11) is a compact locale, since subobjects of A may be identified with open sublocales of E.

More generally, an object A of an arbitrary Grothendieck topos $\mathcal{E}$ is compact iff the topos $\mathcal{E}/A$ is compact (over **Set**) in the sense of C3.2.12; but we shall not make use of any results from Section C3.2 here. We shall need the following stability properties of compactness:

Lemma 3.3.3

(i) *Let A be a compact object of a geometric category. Then, for any cover $f\colon A \twoheadrightarrow B$, B is also compact.*

(ii) *A disjoint union $A \amalg B$ is compact iff both A and B are compact.*

Proof (i) follows from the fact that $f^*\colon \mathrm{Sub}(B) \to \mathrm{Sub}(A)$ preserves unions and is conservative. (ii) follows from the isomorphism $\mathrm{Sub}(A \amalg B) \cong \mathrm{Sub}(A) \times \mathrm{Sub}(B)$. □

The link with 3.3.1 is provided by

Lemma 3.3.4 *Let $(\mathcal{C}, T)$ be a small site, $U \in \mathrm{ob}\,\mathcal{C}$. Then the associated sheaf $l(U)$ of the functor $\mathcal{C}(-, U)$ is compact in $\mathbf{Sh}(\mathcal{C}, T)$ iff U is 'T-compact' in $\mathcal{C}$; i.e. every T-covering sieve on U contains a finite family which generates a T-covering sieve.*

Proof This is immediate from the fact (C2.3.2(c)) that a sieve R on $l(U)$ in $\mathbf{Sh}(\mathcal{C}, T)$ is epimorphic iff the sieve $\{f\colon V \to U \mid l(f) \in R\}$ is T-covering in $\mathcal{C}$. □

For an arbitrary sheaf A on $(\mathcal{C}, T)$, we always have an epimorphic family of morphisms $l(U) \to A$ (as U ranges over all the objects of $\mathcal{C}$); so a necessary condition for compactness of A is the existence of a finite epimorphic family $(f_i \colon l(U_i) \to A \mid 1 \le i \le n)$. If the $l(U_i)$ are all compact, then this condition is also sufficient, by 3.3.3.

However, the presence of finite limits in $\mathcal{C}$ implies an important additional property of the objects $l(U)$. We shall say that an object A of a geometric category is *stable* if, whenever we are given a morphism $f \colon B \to A$ where B is compact, the domain R of the kernel-pair $R \rightrightarrows B$ of f is compact. And we shall call A *coherent* if it is both compact and stable. (Stability is more usually defined by reference to pullbacks of pairs of (not necessarily equal) morphisms from compact objects to A; but, using 3.3.3(ii), it is easy to prove that this definition is equivalent to the simpler one just given.)

In a general topos, stability may be a vacuous condition, because there are 'not enough' compact objects: for example, in the topos $\mathbf{Sh}(\mathbb{R})$ of sheaves on the real line, it follows easily from 3.3.2 that the only compact object is the initial object, and so every object is stable. But in the topos of sheaves on a site where every covering family is finite, 3.3.4 assures us of a plentiful supply of compact objects, and so stability becomes a nontrivial restriction.

Example 3.3.5 Let G be a group, and consider the topos $[G, \mathbf{Set}]$ of (left) G-sets. It is trivial to verify that a G-set A is compact iff it has finitely many G-orbits; in particular, G itself (with G-action via left translations) is a compact G-set. Hence, if A is a stable G-set and $x \in A$, the domain of the kernel-pair of $f \colon G \to A$ (where f is the unique G-equivariant map sending 1 to x) must have finitely many orbits. But it is easy to see that the orbits of this object correspond bijectively to elements of the stabilizer subgroup $G_x = \{g \in G \mid gx = x\}$; so this must be finite. Conversely, if A is any G-set whose point-stabilizers are all finite, it is easy to verify that A is stable.

Lemma 3.3.6

(i) *A subobject of a stable object is stable.*

(ii) *A coproduct of stable objects is stable.*

Proof (i) is easy since if $A' \rightarrowtail A$ is monic then the kernel-pair of any morphism $B \to A'$ coincides with that of the composite $B \to A' \rightarrowtail A$.

(ii) Suppose A_1 and A_2 are stable, and we are given a morphism $f \colon B \to A_1 \amalg A_2$ with B compact. Then B decomposes as a disjoint union $B_1 \amalg B_2$ (where B_i is the pullback of B along the coprojection ν_i), and we have $B \times_{A_1 \amalg A_2} B \cong (B_1 \times_{A_1} B_1) \amalg (B_2 \times_{A_2} B_2)$, which is compact by 3.3.3(ii). This proves the result for finite coproducts; that for infinite coproducts follows, because if $f \colon B \to \coprod_{i \in I} A_i$ is a morphism from a compact object to an infinite coproduct, there can be only finitely many i such that $\nu_i^*(B)$ is nonzero. □

Theorem 3.3.7 *Let* $(\mathcal{C}, T)$ *be a site satisfying (the coherent version of) the hypotheses of* 3.3.1(ii). *Then, for every object* U *of* $\mathcal{C}$, *the object* $l(U)$ *is coherent*

in $\mathbf{Sh}(\mathcal{C},T)$. *If* $(\mathcal{C},T)$ *satisfies the hypotheses of* 3.3.1(iv), *then the converse holds: every coherent object of* $\mathbf{Sh}(\mathcal{C},T)$ *is isomorphic to* $l(U)$ *for some* U.

Proof By 3.3.4, we already know that $l(U)$ is compact, so we have only to prove that it is stable. Suppose given a morphism $A \to l(U)$ with A compact; then by the remarks after 3.3.4 we have a finite epimorphic family $(l(V_i) \to A \mid 1 \leq i \leq n)$. So $A \times_{l(U)} A$ is covered by the pullbacks $l(V_i) \times_{l(U)} l(V_j)$, $1 \leq i, j \leq n$, and it suffices by 3.3.3 to prove that each of these is compact. Since we are not assuming that the coverage T is subcanonical, the composites $l(V_i) \to A \to l(U)$ need not derive from morphisms $V_i \to U$ in $\mathcal{C}$; but, at the cost of replacing each V_i by a finite T-covering family, we may assume that they do. But then since l is cartesian we have $l(V_i) \times_{l(U)} l(V_j) \cong l(V_i \times_U V_j)$, which is compact by 3.3.4.

For the second assertion, let A be an arbitrary coherent object of $\mathbf{Sh}(\mathcal{C},T)$ (where we are now assuming that $(\mathcal{C},T)$ is a pretopos with its coherent coverage). Since A is compact, we can find a finite epimorphic family $(l(U_i) \to A \mid 1 \leq i \leq n)$; and since $l\colon \mathcal{C} \to \mathbf{Sh}(\mathcal{C},T)$ is a coherent functor, we can reduce this to a single epimorphism $l(U) \cong \coprod_{i=1}^{n} l(U_i) \twoheadrightarrow A$, where $U = \coprod_{i=1}^{n} U_i$ in $\mathcal{C}$. Let $R \rightrightarrows l(U)$ be the kernel-pair of $l(U) \twoheadrightarrow A$; then R is compact since A is stable, and so we can find a further epimorphism $l(V) \twoheadrightarrow R$. Now the composite $l(V) \twoheadrightarrow R \rightarrowtail l(U) \times l(U) \cong l(U \times U)$ is in the image of l, since l is full; so its image R is also in the image of l. Finally, since $\mathcal{C}$ has and l preserves coequalizers of equivalence relations, we see that A is in the image of l. □

Theorems 3.3.1 and 3.3.7 together tell us that, up to equivalence, coherent toposes are 'the same thing as' small pretoposes. We may thus fulfil a promise made in Section D3.1:

Corollary 3.3.8 *A coherent theory* $\mathbb{T}$ *is determined up to Morita equivalence by its classifying topos* $\mathbf{Set}[\mathbb{T}]$.

Proof We defined Morita equivalence, for coherent theories, to be equivalence of the pretoposes $\mathcal{P}_{\mathbb{T}}$ generated by their syntactic categories. But by 3.3.7 we can recover $\mathcal{P}_{\mathbb{T}}$ from $\mathbf{Set}[\mathbb{T}]$, as its full subcategory of coherent objects. □

Remark 3.3.9 It also follows immediately from 3.3.7 that, in a coherent topos $\mathcal{E}$, the full subcategory $\mathcal{E}_{\mathrm{coh}}$ of coherent objects is closed under finite limits, finite coproducts and coequalizers of equivalence relations. (In fact it is possible to give a direct proof of this, without invoking the theory of classifying toposes – this is done in [36] and [504]. However, these closure properties, unlike the elementary ones of 3.3.3 and 3.3.6, do not hold in an arbitrary topos.) But $\mathcal{E}_{\mathrm{coh}}$ need not be closed under arbitrary coequalizers, even if it has them. Let $\mathbf{Set}_c$ be the full subcategory of $\mathbf{Set}$ whose objects are all finite or countable sets; then $\mathbf{Set}_c$ is a pretopos, and so we obtain a coherent topos $\mathbf{Sh}(\mathbf{Set}_c, P)$ of sheaves on it for the

coherent coverage. However, the coequalizer diagram

$$\mathbb{N} \underset{1}{\overset{s}{\rightrightarrows}} \mathbb{N} \longrightarrow 1$$

(where $\mathbb{N}$ is the set of natural numbers) is not preserved by the embedding $l \colon \mathbf{Set}_c \to \mathbf{Sh}(\mathbf{Set}_c, P)$. For if it were, then since the coproduct diagram of A2.5.5(i) is also preserved by l, it would follow from 5.1.2 that $l(\mathbb{N})$ would be a natural number object in $\mathbf{Sh}(\mathbf{Set}_c, P)$. But the natural number object in any Grothendieck topos is a countable copower of 1, so it cannot be compact unless the topos is degenerate. Hence this contradicts 3.3.4. (See also C4.2.4, for a close relative of this counterexample.)

Remark 3.3.10 As we remarked earlier, it is possible to give an analogous characterization of the classifying toposes of regular theories, replacing the notion of compactness by that of supercompactness (an object A of a geometric category is called *supercompact* if every covering family with codomain A contains a singleton cover), and coherence by the corresponding notion of supercoherence. In particular, if $\mathcal{E}$ is a topos satisfying the 'regular' version of the conditions of 3.3.1, then the full subcategory of supercoherent objects of $\mathcal{E}$ is a small effective regular category, and $\mathcal{E}$ is equivalent to the category of sheaves on it for the regular coverage; conversely, if $\mathcal{C}$ is any small effective regular category, then the supercoherent objects of $\mathbf{Sh}(\mathcal{C})$ are exactly the representable functors. We omit the proofs, which are very similar to those that we have given in the coherent case.

Using the notion of supercompactness introduced in 3.3.10, we may deduce an interesting syntactic property of regular theories, which will be of use in Section D3.5:

Lemma 3.3.11 *Let $\mathbb{T}$ be a regular theory, and let ϕ and $(\psi_i \mid i \in I)$ be regular formulae over the signature of $\mathbb{T}$ such that the geometric sequent $(\phi \vdash_{\vec{x}} \bigvee_{i \in I} \psi_i)$ is derivable in $\mathbb{T}$ (using geometric logic). Then, for some $i \in I$, the regular sequent $(\phi \vdash_{\vec{x}} \psi_i)$ is also derivable.*

Proof If $(\phi \vdash_{\vec{x}} \bigvee_{i \in I} \psi_i)$ is derivable, then it must be satisfied in the generic $\mathbb{T}$-model in $\mathbf{Set}[\mathbb{T}]$: that is, $l(\{\vec{x} \,.\, \phi\})$ is contained in the union of the $l(\{\vec{x} \,.\, \psi_i\})$ as a subobject of $l(\{\vec{x} \,.\, \top\})$. (Here l, as usual, denotes the Yoneda embedding $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}} \to \mathbf{Set}[\mathbb{T}]$.) But $l(\{\vec{x} \,.\, \phi\})$ is supercompact, and so it must be contained in one of the $l(\{\vec{x} \,.\, \psi_i\})$. Since l is conservative, this means that we have $\{\vec{x} \,.\, \phi\} \leq \{\vec{x} \,.\, \psi_i\}$ in $\mathcal{C}_{\mathbb{T}}^{\mathrm{reg}}$, and hence by (the regular analogue of) 1.4.4(iv) the sequent $(\phi \vdash_{\vec{x}} \psi_i)$ is derivable in $\mathbb{T}$. □

There is of course an analogue of 3.3.11 for coherent theories, in which an arbitrary disjunction is reduced to a finite one; we leave the detailed formulation of this result to the reader. We should also mention that there is an alternative, more 'classically model-theoretic', proof of 3.3.11 using the classical completeness

theorem 1.5.4(ii) and the fact (which we noted after 2.4.3) that the class of models of a regular theory is closed under products: for if the conclusion of 3.3.11 fails, then for each i we can find a $\mathbb{T}$-model M_i in **Set** containing a tuple of elements which belongs to the interpretation of $\vec{x}\,.\,\phi$ but not of $\vec{x}\,.\,\psi_i$, and then the product $\prod_{i \in I} M_i$ is easily seen to contain a tuple which belongs to $[\![\vec{x}\,.\,\phi]\!]$ but not to $[\![\vec{x}\,.\bigvee_{i \in I} \psi_i]\!]$. (Using this argument, we may also prove a converse to the remark after 2.4.3: if $\mathbb{T}$ is a coherent theory over a signature Σ, such that $\mathbb{T}$-Mod(**Set**) is closed under arbitrary products in Σ-Str(**Set**), then $\mathbb{T}$ is equivalent to a regular theory over Σ.)

Remark 3.3.12 There is a close connection between the notion of coherence for objects of a topos, studied in this section, and the notion of finite presentability studied in Section D2.3. It is easy to see that a finitely presentable object A of a geometric category is compact; for if A is the union of a directed family of subobjects A_i, then the identity morphism $A \to A$ must factor through some $A_i \rightarrowtail A$. Conversely, if $(\mathcal{C}, T)$ is a site such that T is generated by finite covering families, then it is easy to see that $\mathbf{Sh}(\mathcal{C}, T)$ is closed under filtered colimits in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, and hence that the sheaves $l(U)$ are finitely presentable in $\mathbf{Sh}(\mathcal{C}, T)$; so it follows from 3.3.7 that every coherent object in a coherent topos is finitely presentable. However, neither of these implications is reversible. In any cocomplete category, the finitely presentable objects are closed under arbitrary finite colimits, so that (for example) in the topos mentioned in 3.3.9, the coequalizer of $l(\mathbb{N}) \rightrightarrows l(\mathbb{N})$ is finitely presentable but not coherent. And if X is any compact space which is not strongly compact in the sense of Section C3.4 (for example, the space of C3.4.1(a)), then the terminal object of $\mathbf{Sh}(X)$ is compact (by 3.3.2) but not finitely presentable.

We note in particular that a coherent topos is a locally finitely presentable category (this could of course have been deduced from 2.1.4(h) and 2.3.7). Once again, the converse is false: the topos $\mathbf{Set}[\mathbb{K}]$ of 3.2.10 provides a counterexample, as does the topos $[G, \mathbf{Set}]$ for any infinite group G (cf. 3.3.5).

The classical completeness theorem for coherent theories (1.5.10) may be translated into an important property of coherent toposes:

Proposition 3.3.13 *A coherent topos $\mathcal{E}$ has enough points; that is, the inverse image functors $\mathcal{E} \to$ **Set** are jointly conservative.*

Proof Suppose $\mathcal{E} \simeq \mathbf{Set}[\mathbb{T}]$, and let $(M_i \mid i \in I)$ be a set of $\mathbb{T}$-models in **Set** which are sufficient to determine provability in $\mathbb{T}$. We may regard these as a single $\mathbb{T}$-model M in $\mathbf{Set}/I$, and we need to show that its classifying map $f\colon \mathbf{Set}/I \to \mathbf{Set}[\mathbb{T}]$ is a surjection. We do this by the method of 3.2.6(ii): let $\mathcal{C}$ be the class of all objects A of $\mathbf{Set}[\mathbb{T}]$ such that f^* preserves properness of subobjects of A. Since M is conservative for coherent sequents, we see that $\mathcal{C}$ includes (the images in $\mathbf{Set}[\mathbb{T}]$ of) all the objects of the syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$; but these form a separating set for $\mathbf{Set}[\mathbb{T}]$, and $\mathcal{C}$ is closed under coproducts and quotients, so it contains all the objects of $\mathbf{Set}[\mathbb{T}]$. $\square$

Proposition 3.3.13 may also be proved by 'purely topos-theoretic' methods, without reference to coherent theories or classifying toposes; this is done in [36] and [504]. However, inspection of the proof shows that it involves precisely the same sort of iterative construction that went into the proof of the completeness theorems in Section D1.5. This is not surprising, since the two are formally equivalent: the deduction of 1.5.10(ii) from the statement of 3.3.13 is even easier than the proof of 3.3.13, given that validity in the generic model of $\mathbb{T}$ is equivalent to provability in $\mathbb{T}$.

Remark 3.3.14 If $\mathbb{T}$ is a coherent propositional theory, then $\mathbf{Set}[\mathbb{T}]$ is localic over $\mathbf{Set}$ (by B4.2.12, or alternatively because $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ is a preorder, as we observed in 1.4.14); so by 3.3.13 it is actually spatial, i.e. of the form $\mathbf{Sh}(X)$ for a topological space X. The spaces X which occur in this context are easy to describe: they are exactly the *coherent spaces*, i.e. those whose topology has a basis which is closed under finite intersections (including the empty intersection X) and consists of open sets which are compact. If X is a coherent space, then the poset L of all compact open subsets of X is a distributive lattice, and X is homeomorphic to the *spectrum* spec L, i.e. the space whose points are the prime filters of L, with topology having as a basis the sets $\{F \in \operatorname{spec} L \mid U \in F\}$, $U \in L$; the Comparison Lemma (C2.2.3) then ensures that $\mathbf{Sh}(X)$ is equivalent to the topos of sheaves on L for the coherent coverage. So the fact that every localic coherent topos is of this form is another reflection of the fact, observed in 1.4.14, that every coherent propositional theory is Morita-equivalent to the theory of prime filters of a distributive lattice.

In connection with the foregoing remark, we note

Lemma 3.3.15 *The localic reflection functor* $\mathfrak{BTop}/\mathbf{Set} \to \mathfrak{LTop}/\mathbf{Set}$ *preserves coherent toposes.*

Proof Suppose $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, T)$ is a coherent topos, where $\mathcal{C}$ is a small pretopos and T is its coherent coverage. Then $(\mathcal{C}, T)$ satisfies the hypotheses of C2.3.22, so by that result we may identify the localic reflection of $\mathcal{E}$ with $\mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), T')$ where T' is the coverage induced on $\mathrm{Sub}_{\mathcal{C}}(1)$ by T. But $\mathrm{Sub}_{\mathcal{C}}(1)$ is a distributive lattice by A1.4.2, and T' may easily be identified with its coherent coverage; so $\mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), T')$ is a coherent topos. □

Lemma 3.3.15 may alternatively be proved starting from a site for $\mathcal{E}$ which merely satisfies the hypotheses of 3.3.1(ii). In this case, however, we have to use the more complicated description of the localic reflection provided by C2.3.20.

We conclude this section by noting two further closure properties of the class of coherent toposes:

Lemma 3.3.16

(i) *If $\mathcal{E}$ is a coherent topos, then the slice category $\mathcal{E}/B$ is a coherent topos iff B is a coherent object of $\mathcal{E}$.*

(ii) *If $\mathcal{E}$ and $\mathcal{F}$ are Grothendieck toposes, the product category $\mathcal{E} \times \mathcal{F}$ is a coherent topos iff both $\mathcal{E}$ and $\mathcal{F}$ are.*

Proof (i) It is easily verified that an object $f: A \to B$ of $\mathcal{E}/B$ is compact (resp. coherent) iff A is compact (resp. coherent) in $\mathcal{E}$. So if B is coherent, the subcategory $(\mathcal{E}/B)_{\mathrm{coh}}$ of coherent objects of $\mathcal{E}/B$ is simply $\mathcal{E}_{\mathrm{coh}}/B$; hence it is a pretopos, and its objects form a separating family for $\mathcal{E}/B$. The converse follows from the fact that the terminal object of a coherent topos is a coherent object.

(ii) By 3.3.3(ii) and 3.3.6(i), a complemented subterminal object in a coherent topos is coherent; so the left-to-right implication follows from (i). For the converse, it is again easy to verify that an object (A, B) of $\mathcal{E} \times \mathcal{F}$ is compact (resp. coherent) iff A is compact (resp. coherent) in $\mathcal{E}$ and B is compact (resp. coherent) in $\mathcal{F}$. So if $\mathcal{E}$ and $\mathcal{F}$ are coherent, the coherent objects of $\mathcal{E} \times \mathcal{F}$ form a pretopos and a separating family. □

Remark 3.3.17 An alternative 'syntactic' proof of the last part of 3.3.16 can be given as follows: let $\mathbb{T}_1$ and $\mathbb{T}_2$ be the coherent theories classified by $\mathcal{E}$ and $\mathcal{F}$ respectively. Since $\mathcal{E} \times \mathcal{F}$ is the coproduct of $\mathcal{E}$ and $\mathcal{F}$ in $\mathfrak{Top}$, and since such coproducts are stable under pullback, a geometric morphism $\mathcal{H} \to \mathcal{E} \times \mathcal{F}$ corresponds to a complementary pair of subterminal objects (U_1, U_2) in $\mathcal{H}$ together with a $\mathbb{T}_1$-model in $\mathcal{H}/U_1$ and a $\mathbb{T}_2$-model in $\mathcal{H}/U_2$. Now it is easy to construct a coherent theory $\mathbb{T}_3$ whose models correspond to the above data: specifically, if Σ_1 and Σ_2 are the signatures of $\mathbb{T}_1$ and $\mathbb{T}_2$, we take a signature Σ_3 which is the disjoint union of these two together with two new sorts U_1 and U_2, and new function symbols $f_A: A \to U_1$ (resp. $g_A: A \to U_2$) for each sort A of Σ_1 (resp. Σ_2). As axioms, we take the axioms of $\mathbb{T}_1$ (with a free variable of sort U_1 added to the context, if it happens to be empty), the axioms of $\mathbb{T}_2$ (similarly modified), and the sequents $(\top \vdash_{x_i, y_i} (x_i = y_i))$ $(i = 1, 2)$,

$$(\top \vdash_{[]} ((\exists x_1)\top \vee (\exists x_2)\top)) \text{ , and}$$

$$(((\exists x_1)\top \wedge (\exists x_2)\top) \vdash_{[]} \bot) \text{ ,}$$

where x_i and y_i are variables of sort U_i $(i = 1, 2)$. It is now clear that if M is a $\mathbb{T}_3$-model in a topos $\mathcal{H}$, then MU_1 and MU_2 are complementary subterminal objects, and the remaining data yield models of $\mathbb{T}_1$ and $\mathbb{T}_2$ in the corresponding slice categories.

At least in the case when B is (the image in $\mathcal{E}$ of) an object of the syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$, where $\mathbb{T}$ is the coherent theory classified by $\mathcal{E}$, a similar proof may be given for the first part of 3.3.16. For if B is the object $\{\vec{x}.\phi\}$, then $\mathcal{E}/B$ classifies the theory obtained from $\mathbb{T}$ by adding a string of new constants $\vec{c}$ of the same length and type as $\vec{x}$, together with the new axiom $(\top \vdash_{[]} \phi[\vec{c}/\vec{x}])$.

We note that a regular topos is connected and locally connected over **Set**, by C3.3.11(b). It follows that the second part of 3.3.16 does not hold for regular toposes, though the first part does (by either of the two proofs given in

3.3.16 and 3.3.17), and so does the regular analogue of 3.3.15. Also, although 3.3.16(ii) does not hold for regular toposes, it does hold for the classifying toposes of disjunctive theories, as may readily be seen from the proof given in 3.3.17.

Suggestions for further reading: Artin et al. [36], Makkai [777].

D3.4 Boolean classifying toposes

Given the 'equivalence' between Grothendieck toposes and geometric theories established in Section D3.1, it is naturally of interest to ask how familiar properties of toposes relate to familiar (syntactic or model-theoretic) properties of the theories which they classify. In this section, we shall be concerned with the question of what can be said about a theory whose classifying topos is Boolean; we shall be mainly interested in coherent theories, for which this question has a particularly satisfying answer, first described by A. Blass and A. Ščedrov [133]. We shall briefly discuss infinitary geometric theories whose classifying toposes are Boolean at the end of the section.

To begin with, we need to investigate what can be said about Boolean toposes which are coherent, in the sense described in the last section. We start with a particular case: when is the topos $\mathbf{Cont}(G)$ of continuous G-sets for a topological group G (cf. A2.1.6) coherent? We recall that, given subgroups H and K of a group G, a *double coset* of the pair (H, K) is a subset of G of the form $HgK = \{hgk \mid h \in H, k \in K\}$ for some $g \in G$. Just as for ordinary cosets, it is easy to prove that the double cosets of a given pair of subgroups form a partition of G (that is, if HgK and $Hg'K$ overlap, they must be identical). We shall be particularly concerned with the case $H = K$; we shall say that a subgroup H has *finite bi-index* in G if the number of distinct double cosets HgH, $g \in G$, is finite. And we shall say that a topological group G is *coherent* if every open subgroup of G has finite bi-index. Note that for a normal subgroup H, double cosets HgH coincide with ordinary cosets gH; hence, for a pro-discrete group G (one expressible as an inverse limit of discrete groups), coherence is equivalent to profiniteness. However, for non-normal subgroups, having finite bi-index is a weaker condition than having finite index:

Example 3.4.1 Let G be the group of all permutations of $\mathbb{N}$, and let H be the pointwise stabilizer of a finite subset $S \subseteq \mathbb{N}$, i.e. the set of all $g \in G$ satisfying $gn = n$ for all $n \in S$. It is clear that, except in the case $S = \emptyset$, H has infinite index in G. However, it has finite bi-index: for if two elements g and g' of G satisfy $\{n \in S \mid gn \in S\} = \{n \in S \mid g'n \in S\}$ and the restrictions of g and g' to this subset are equal, then it is easy to see that we can obtain g' from g by pre- and post-multiplying by elements of H, and hence $HgH = Hg'H$. Thus the number of double cosets of H is $\sum_{k=0}^{n}\binom{n}{k}^2 k!$, where n is the number of elements of S; although this number increases quite rapidly with n, it is undeniably finite. Hence if we equip G with the topology of pointwise convergence, in which the

pointwise stabilizers of finite subsets of $\mathbb{N}$ form a base of neighbourhoods of the identity element, then it becomes a coherent topological group.

Lemma 3.4.2 *Let G be a topological group. Then the topos* $\mathbf{Cont}(G)$ *is coherent iff G is coherent in the sense defined above.*

Proof First we note that, as in the topos $[G, \mathbf{Set}]$ of all G-sets (3.3.5), the compact objects of $\mathbf{Cont}(G)$ are exactly the continuous G-sets with finitely many orbits. If $\mathbf{Cont}(G)$ is coherent, then its terminal object 1 must be stable, i.e. $A \times A$ must be compact for every compact A; so in particular the product $G : H \times G : H$, where H is an open subgroup of G (and $G : H$ denotes the set of left cosets of H, with G acting by left translation), must have finitely many G-orbits. But in fact this condition (for all open subgroups of G) is sufficient; for it implies that $G : H \times G : K$ is compact for any two open subgroups H and K (since $H \cap K$ is open if H and K are, and $G : H$ and $G : K$ are both epimorphic images of $G : (H \cap K)$). Hence the square of any compact object, being a finite coproduct of objects of this form, is compact, and so (since any subobject of a compact object is compact) it follows that every object is stable. Given this, it is easy to see that the compact (equivalently, coherent) objects of $\mathbf{Cont}(G)$ form a pretopos, and that they form a site of definition for $\mathbf{Cont}(G)$ when equipped with the coherent coverage.

Thus we are reduced to determining when $G : H \times G : H$ has finitely many orbits. But (xH, yH) is in the same orbit as $(x'H, y'H)$ iff there exist elements $g \in G$ and $h, k \in H$ such that $gxh = x'$ and $gyk = y'$, iff the double cosets $Hx^{-1}yH$ and $Hx'^{-1}y'H$ are equal. So a necessary and sufficient condition for coherence of $\mathbf{Cont}(G)$ is that each open subgroup H should have finite bi-index. □

Theorem 3.4.3 *For a Grothendieck topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ is Boolean and coherent.*

(ii) *$\mathcal{E}$ is atomic (over* $\mathbf{Set}$*) and coherent.*

(iii) *$\mathcal{E}$ is equivalent to a finite coproduct of toposes of the form* $\mathbf{Cont}(G_i)$*, where the G_i are coherent topological groups.*

Proof (ii) $\Rightarrow$ (i) is trivial since atomic toposes are Boolean (C3.5.2), and (iii) $\Rightarrow$ (ii) follows from 3.4.2 and the fact that a finite coproduct of coherent toposes is coherent (3.3.16(ii)).

(i) $\Rightarrow$ (ii): Suppose $\mathcal{E}$ is a Boolean coherent topos. By 3.3.3(ii) and 3.3.6(i), any subobject of a coherent object A of $\mathcal{E}$ is coherent; moreover, since the lattice $\mathrm{Sub}(A)$ is Boolean and its top element is compact, it must actually be finite and hence atomic. So every object of $\mathcal{E}_{\mathrm{coh}}$ is a finite coproduct of atoms (and hence covered by atoms in the coherent coverage), and so we may replace $\mathcal{E}_{\mathrm{coh}}$, as a site of definition for $\mathcal{E}$, by its full subcategory $\mathcal{E}_{\mathrm{at}}$ consisting of atoms. And this category, with the coverage induced by the coherent coverage on $\mathcal{E}_{\mathrm{coh}}$, is easily seen to satisfy the hypotheses of C3.5.8; so $\mathcal{E}$ is atomic.

(ii) $\Rightarrow$ (iii): If we add to (ii) the hypothesis that $\mathcal{E}$ is connected (i.e. that its terminal object is an atom), then (since it must have a point, by 3.3.13) it satisfies the hypotheses of C5.2.13, and so must be of the form $\mathbf{Cont}(G)$; moreover, by 3.4.2, the topological group G must be coherent. In general, the terminal object 1 is a finite coproduct $\coprod_{i=1}^{n} A_i$ of atoms (since it is coherent), and so $\mathcal{E}$ is a finite coproduct of connected atomic toposes $\mathcal{E}/A_i$, which are coherent by 3.3.16(ii). So the result follows from the particular case already discussed. $\square$

We note that the equivalence of (i) and (ii) in 3.4.3 could alternatively be deduced from 3.3.13 and C3.5.2.

We now turn to the question of what can be said about a coherent theory $\mathbb{T}$ if its classifying topos $\mathbf{Set}[\mathbb{T}]$ satisfies the conditions of 3.4.3. First we note

Scholium 3.4.4 *Let $\mathbb{T}$ be a coherent theory over a signature Σ, such that $\mathbf{Set}[\mathbb{T}]$ is Boolean. Then, for any context $\vec{x}$ over Σ, there are only finitely many $\mathbb{T}$-provable-equivalence classes of coherent formulae in the context $\vec{x}$.*

Proof Such provable-equivalence classes correspond to (isomorphism classes of) subobjects of $\{\vec{x}\,.\top\}$ in $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$, by 1.4.4(iv). But the canonical functor $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}} \to \mathbf{Set}[\mathbb{T}]$ is conservative, and we have seen in the proof of 3.4.3(i) $\Rightarrow$ (ii) that any coherent object of $\mathbf{Set}[\mathbb{T}]$ (in particular, any object in the image of this functor) has finite subobject lattice. $\square$

In order to proceed further, we shall find it convenient to work not with the whole syntactic category $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ but with its full subcategory $\mathcal{C}'$ of *consistent objects*, i.e. those objects $\{\vec{x}\,.\phi\}$ for which the sequent $(\phi \vdash_{\vec{x}} \bot)$ is not provable in $\mathbb{T}$. The point of this is that the inconsistent objects are precisely those which are covered by the empty family in the coherent coverage P on $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$; so if we delete them we obtain (as in C2.2.4(*e*)) a site $(\mathcal{C}', P')$ which gives rise to the same topos, but in which every covering family is nonempty – equivalently, $\mathbf{Sh}(\mathcal{C}', P')$ is a dense subtopos of $[\mathcal{C}'^{\mathrm{op}}, \mathbf{Set}]$. So, if the classifying topos of $\mathbb{T}$ is Boolean, we may identify P' with the coverage corresponding to the double-negation local operator on $[\mathcal{C}'^{\mathrm{op}}, \mathbf{Set}]$ (cf. A4.5.21); in other words, a sieve R on an object A of $\mathcal{C}'$ is P'-covering iff it is 'stably nonempty', i.e. its pullback along any $f\colon B \to A$ in $\mathcal{C}'$ is nonempty.

Lemma 3.4.5 *Suppose $\mathbb{T}$ is a coherent theory such that $\mathbf{Set}[\mathbb{T}]$ is Boolean. Then, for every coherent formula-in-context $\vec{x}\,.\phi$ over the signature of $\mathbb{T}$, there exists a coherent formula $\vec{x}\,.\psi$ in the same context such that the sequents $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$ and $(\top \vdash_{\vec{x}} (\phi \vee \psi))$ are provable in $\mathbb{T}$.*

Proof By 3.4.4, there are (up to $\mathbb{T}$-provable equivalence) only finitely many coherent formulae ψ for which the first sequent holds, so (by taking their disjunction) we may find a single such ψ which is entailed by all the others. Suppose this ψ does not satisfy the second sequent; then the sieve on the object $\{\vec{x}\,.\top\}$ of $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ generated by the canonical monomorphisms $\{\vec{x}\,.\phi\} \rightarrowtail \{\vec{x}\,.\top\}$ and

$\{\vec{x}.\psi\} \rightarrowtail \{\vec{x}.\top\}$ is not covering for the coherent coverage. (We shall assume that the domains of both these morphisms are consistent; if either of them is inconsistent, the argument is similar but easier.)

From the description given above of the coverage P', we deduce that there exists a consistent object $\{\vec{y}.\chi\}$ and a morphism

$$\{\vec{y}.\chi\} \xrightarrow{[\theta]} \{\vec{x}.\top\}$$

such that no morphism with consistent domain factors through both $[\theta]$ and one of the two monomorphisms above. In particular, this implies that $\{\vec{x},\vec{y}.(\theta \wedge \phi)\}$ and $\{\vec{x},\vec{y}.(\theta \wedge \psi)\}$ are both inconsistent. But the first of these implies that $(((\exists\vec{y})\theta \wedge \phi) \vdash_{\vec{x}} \bot)$ is provable in $\mathbb{T}$, so by the construction of ψ the sequent $((\exists\vec{y})\theta \vdash_{\vec{x}} \psi)$ is provable, and hence $(\theta \vdash_{\vec{x},\vec{y}} (\theta \wedge \psi))$ is provable. So $\{\vec{x},\vec{y}.\theta\}$ is inconsistent; but since θ represents a morphism of $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ with domain $\{\vec{y}.\chi\}$, the sequent $(\chi \vdash_{\vec{y}} (\exists\vec{x})\theta)$ is provable, contradicting the consistency of $\{\vec{y}.\chi\}$. □

We are now ready for the main result of this section.

Theorem 3.4.6 *Let $\mathbb{T}$ be a coherent theory over a signature Σ. Then the following conditions are equivalent:*

(i) *The classifying topos* **Set**$[\mathbb{T}]$ *is Boolean.*

(ii) *$\mathbb{T}$ satisfies the following two conditions:*

  (a) *for every first-order formula-in-context $\vec{x}.\phi$ over Σ, there exists a coherent formula $\vec{x}.\psi$ in the same context, such that the sequents $(\phi \dashv\vdash_{\vec{x}} \psi)$ are provable in $\mathbb{T}$ using classical logic; and*

  (b) *in any context over Σ, there are only finitely many first-order formulae up to classical $\mathbb{T}$-provable equivalence.*

Proof First suppose **Set**$[\mathbb{T}]$ is Boolean. To prove that (ii)(a) holds, we proceed by induction over the structure of ϕ. Since provable equivalence is stable under compounding, we need only consider the cases when ϕ is obtained by negation, implication or universal quantification from formulae for which the result is known to be true. Lemma 3.4.5 takes care of negation; for the other two cases, we use the fact that the sequents $((\theta \Rightarrow \chi) \dashv\vdash (\neg\theta \vee \chi))$ and $((\forall x)\theta \dashv\vdash \neg(\exists x)\neg\theta)$ are classically provable. Condition (ii)(b) is now immediate from (ii)(a) and 3.4.4.

Conversely, suppose $\mathbb{T}$ satisfies the conditions in (ii). By the classical completeness theorem 1.5.10 (or by 3.1.16, if you prefer), the relations of classical $\mathbb{T}$-provable equivalence and constructive $\mathbb{T}$-provable equivalence coincide for coherent formulae, and so 1.4.4(iv) implies that the subobject lattice of any object of $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ is a finite Boolean algebra, and in particular atomic. As before, let $\mathcal{C}'$ be the full subcategory of $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$ obtained by deleting the inconsistent objects, and let P' be the coverage induced on it by the coherent coverage on $\mathcal{C}_{\mathbb{T}}^{\mathrm{coh}}$; we shall show that P' coincides with the coverage N consisting of stably nonempty

sieves, so that $\mathbf{Set}[\mathbb{T}] \simeq \mathbf{Sh}(\mathcal{C}', P') \simeq \mathbf{sh}_{\neg\neg}[\mathcal{C}'^{\mathrm{op}}, \mathbf{Set}]$. The inclusion $P' \subseteq N$ is clear, since every object of $\mathcal{C}'$ is consistent. So suppose R is an N-covering sieve on an object $\{\vec{x}.\phi\}$ of $\mathcal{C}'$; then, for each atom $\{\vec{x}.\psi\}$ of $\mathrm{Sub}(\{\vec{x}.\phi\})$, there exists a morphism $f\colon \{\vec{y}.\chi\} \to \{\vec{x}.\phi\}$ in R which factors through $\{\vec{x}.\psi\} \rightarrowtail \{\vec{x}.\phi\}$. But $\{\vec{y}.\chi\}$ is consistent and $\{\vec{x}.\psi\}$ has no proper subobjects in $\mathcal{C}'$, so the image of this morphism must be precisely $\{\vec{x}.\psi\}$. Thus if we choose one such morphism for each atom, we obtain a finite family of morphisms in R, the union of whose images is the whole of $\{\vec{x}.\phi\}$. So R is P'-covering. □

Condition (ii)(*b*) of 3.4.6 is well-known to model-theorists as 'Ryll-Nardzewski's criterion for $\aleph_0$-categoricity'. Specifically, we have the following result:

Proposition 3.4.7 *Let $\mathbb{T}$ be a (consistent) complete first-order theory over a signature Σ. If there are only finitely many classical-$\mathbb{T}$-provable-equivalence classes of first-order formulae in any context over Σ, then $\mathbb{T}$ is $\aleph_0$-categorical; that is, it has only one isomorphism class of (finite or) countable models in* **Set**. *If the signature Σ is countable (i.e. has countably many sorts, function symbols and relation symbols), then the converse holds.* □

The proof of 3.4.7 is beyond the scope of this book; see [446, 7.3.1]. We recall that '$\mathbb{T}$ is complete' means that every sentence over Σ is $\mathbb{T}$-provably equivalent to $\top$ or $\bot$ (but not both); in our context, this corresponds to saying that the classifying topos $\mathbf{Set}[\mathbb{T}]$ is two-valued (equivalently, connected). Ryll-Nardzewski's criterion is normally stated only for single-sorted theories; but in fact it works equally well for many-sorted ones, provided that the cardinality of a **Set**-model M of a many-sorted theory is understood to be the cardinality of the disjoint union $\coprod_{A \in \Sigma\text{-Sort}} MA$.

Corollary 3.4.8 *Let $\mathbb{T}$ be a coherent theory such that $\mathbf{Set}[\mathbb{T}]$ is Boolean. Then $\mathbb{T}$ has finitely many completions $\mathbb{T}_1, \ldots, \mathbb{T}_n$, all of which are coherent and $\aleph_0$-categorical; and in fact $\mathbf{Set}[\mathbb{T}]$ is the coproduct of the $\mathbf{Set}[\mathbb{T}_i]$ in $\mathfrak{BTop}/\mathbf{Set}$. Moreover, each $\mathbf{Set}[\mathbb{T}_i]$ is equivalent to $\mathbf{Cont}(G_i)$, where G_i is the automorphism group of the unique countable model M_i of $\mathbb{T}_i$, equipped with the topology in which the pointwise stabilizers of finite subsets of M_i form a base of neighbourhoods of the identity.*

Proof The completions of a first-order theory $\mathbb{T}$ correspond to ultrafilters in the (classical) Lindenbaum algebra of $\mathbb{T}$, i.e. the Boolean algebra of classical-$\mathbb{T}$-provable-equivalence classes of sentences over the signature of $\mathbb{T}$. But 3.4.6 tells us that this algebra is finite, so there are only finitely many such ultrafilters; moreover, they are all principal, and generated by the atoms of the Lindenbaum algebra. If ϕ_i is one such atom, the corresponding completion $\mathbb{T}_i$ is obtained from $\mathbb{T}$ by adding the axiom $(\top \vdash_{[]} \phi_i)$; its classifying topos is the slice category $\mathbf{Set}[\mathbb{T}]/[\![[\,].\phi_i]\!]_G$, where G is the generic $\mathbb{T}$-model, and hence it is still Boolean. So all the completions of $\mathbb{T}$ satisfy the Ryll-Nardzewski criterion. Finally, since

the $\mathbf{Set}[\mathbb{T}_i]$ are connected as well as Boolean, they are of the form $\mathbf{Cont}(G_i)$ by C5.2.14(c), where G_i may be taken to be the localic automorphism group of any point of $\mathbf{Set}[\mathbb{T}_i]$. But if we take the point corresponding to the unique countable model, we obtain the group G_i described in the statement. □

Condition (a) of 3.4.6(ii) is perhaps less familiar to classical model-theorists than (b); however, we note that it is satisfied by the Morleyization (as defined in 1.5.13) of any first-order theory $\mathbb{T}$. (In 1.5.13 we showed that every first-order formula over the original signature Σ of $\mathbb{T}$ is classically provably equivalent, in its Morleyization $\mathbb{T}'$, to a coherent formula over the enlarged signature Σ' of the latter; but it is easy to see that the same is true for any first-order formula over Σ', since we can reduce it to a first-order formula over Σ by replacing each instance of one of the new relation symbols by the corresponding instance of the first-order formula over Σ which it 'represents'.) It is also clear that the Morleyization of $\mathbb{T}$ satisfies condition (b) iff $\mathbb{T}$ itself does.

We may also decompose condition (a) as the conjunction of two simpler conditions (a_1) and (a_2):

(a_1) *the negation of every atomic formula over Σ is (classically) $\mathbb{T}$-provably equivalent to a coherent formula; and*

(a_2) *every first-order formula over Σ is classically $\mathbb{T}$-provably equivalent to an existential formula (i.e. one of the form $(\exists\vec{x})\phi$ where ϕ is quantifier-free and $\vec{x}$ is a (possibly empty) string of variables).*

For both these conditions are clearly implied by (a); conversely, if they both hold, we may replace an arbitrary first-order formula by a $\mathbb{T}$-provably equivalent existential formula, then replace formulae of the form $(\phi \Rightarrow \psi)$ by $(\neg\phi \vee \psi)$, pull negations inside conjunctions and disjunctions using the De Morgan laws (cf. 4.6.1), and finally replace any negated atomic formulae in the resulting formula by their coherent equivalents. Now we have

Proposition 3.4.9 *Let $\mathbb{T}$ be a coherent theory over a signature Σ. Then*

(i) *$\mathbb{T}$ satisfies condition (a_1) iff every homomorphism of $\mathbb{T}$-models in* **Set** *is an embedding (cf.* 1.2.10*).*

(ii) *$\mathbb{T}$ satisfies condition (a_2) iff every embedding of $\mathbb{T}$-models in* **Set** *is an elementary embedding.*

(iii) *$\mathbb{T}$ satisfies condition (a) of* 3.4.6(ii) *iff every homomorphism of $\mathbb{T}$-models in* **Set** *is an elementary embedding.*

Proof Part (ii) of this proposition is familiar to classical model-theorists, who call a first-order theory *model-complete* if it satisfies these conditions. We shall prove only part (i) in detail, and refer the reader to [446, 8.3.1], for the proof of (ii) (which is in any case rather similar to that of (i)); part (iii) is of course immediate from (i) and (ii).

First suppose $\mathbb{T}$ satisfies (a_1). Then, for any atomic formula-in-context $\vec{x}.\phi$ and any homomorphism $h\colon M \to N$ of $\mathbb{T}$-models in **Set**, we have a commutative diagram

$$\begin{tikzcd}
\llbracket \vec{x}.\phi \rrbracket_M \arrow[r, tail] \arrow[d] & M(A_1,\ldots,A_n) \arrow[d, "h_{A_1}\times\cdots\times h_{A_n}"] & \llbracket \vec{x}.\neg\phi \rrbracket_M \arrow[l, tail] \arrow[d] \\
\llbracket \vec{x}.\phi \rrbracket_N \arrow[r, tail] & N(A_1,\ldots,A_n) & \llbracket \vec{x}.\neg\phi \rrbracket_N \arrow[l, tail]
\end{tikzcd}$$

(where $A_1,\ldots,A_n$ is the sequence of sorts of the variables in $\vec{x}$), by 1.2.9. Since the pairs of subobjects in each row are complementary, it follows that the left-hand square must be a pullback; but the latter condition, for all atomic ϕ, is the definition of what it means for h to be an embedding.

Conversely, suppose every homomorphism is an embedding. Let $\vec{x}.\phi$ be an atomic formula-in-context; we shall find it convenient to extend our signature Σ to Σ' by adding a string of new constants $\vec{c}$ of the same length and type as the context $\vec{x}$. Let Γ be the set of all coherent formulae γ in the context $\vec{x}$ such that $((\gamma \wedge \phi) \vdash_{\vec{x}} \bot)$ is provable in $\mathbb{T}$; we need to show that $(\top \vdash_{\vec{x}} (\gamma \vee \phi))$ is provable for some $\gamma \in \Gamma$. Let $\mathbb{T}'$ be the theory over Σ' obtained by adding

$$\{(\gamma[\vec{c}/\vec{x}] \vdash_{[]} \bot) \mid \gamma \in \Gamma\}$$

to the axioms of $\mathbb{T}$; suppose $\mathbb{T}'$ has a model M in **Set**. Let Σ'' be the signature obtained from Σ' by adding a new constant of sort A for each element of MA, for each sort A of Σ, let X be the set of all atomic sentences χ over Σ'' such that $(\top \vdash_{[]} \chi)$ is satisfied in M, and let $\mathbb{T}''$ be the theory obtained by adding all these sequents to the axioms of $\mathbb{T}$ (*not* $\mathbb{T}'$). (It is then easy to see that a $\mathbb{T}''$-model is essentially the same thing as a $\mathbb{T}$-model N equipped with a homomorphism $h\colon M \to N$.) For each $\chi \in \mathrm{X}$, let $\overline{\chi}$ be the formula over Σ obtained from it on replacing the constants in $\vec{c}$ by the corresponding variables in $\vec{x}$, and any other new constants which appear by a string of new variables $\vec{y}$. Then, for each finite subset $\{\chi_1,\ldots,\chi_n\}$ of X, the coherent formula-in-context

$$\vec{x}.(\exists \vec{y})(\overline{\chi_1} \wedge \cdots \wedge \overline{\chi_n})$$

is not in Γ, since it becomes true in M when we substitute $\vec{c}$ for $\vec{x}$. So

$$((\chi_1 \wedge \cdots \wedge \chi_n \wedge \phi[\vec{c}/\vec{x}]) \vdash_{[]} \bot)$$

is not provable in $\mathbb{T}$; hence the theory obtained from $\mathbb{T}''$ by adding $(\top \vdash_{[]} \phi[\vec{c}/\vec{x}])$ is consistent, since each finite subset of it is consistent, and by completeness it has a model $h\colon M \to N$. But h is an embedding, and $\phi[\vec{c}/\vec{x}]$ is satisfied in N, so it must be satisfied in M. Since this is true for any model M of $\mathbb{T}'$, we deduce by classical completeness that $(\top \vdash_{[]} \phi[\vec{c}/\vec{x}])$ must be provable in it; hence it is deducible from some finite subset of the axioms of $\mathbb{T}'$. If $\gamma_1,\ldots,\gamma_m$ are the

members of Γ such that the sequents $(\gamma_i[\vec{c}/\vec{x}] \vdash_{[]} \bot)$ occur among these axioms, then it is easy to see that the sequent

$$(\top \vdash_{\vec{x}} (\gamma_1 \vee \cdots \vee \gamma_m \vee \phi))$$

must be satisfied in every $\mathbb{T}$-model, and hence provable in $\mathbb{T}$. But $(\gamma_1 \vee \cdots \vee \gamma_m)$ is in Γ; so it is the coherent formula we seek. $\square$

We remark that the conditions (a_1) and (a_2) are independent, even for theories which satisfy 3.4.6(ii)(b). Of course, any propositional theory satisfies (a_2) (indeed, any embedding between models of a propositional theory is an isomorphism), and it satisfies (b) provided its signature has only finitely many relation symbols; but the theory of prime filters of a non-Boolean distributive lattice fails to satisfy (a_1). We shall give an example of a theory which satisfies (a_1) but not (a_2) after 3.4.11 below.

We next discuss some important examples of coherent theories whose classifying toposes are Boolean (and connected).

Example 3.4.10 The theory $\mathbb{D}_\infty$ of infinite decidable objects is obtained from the theory $\mathbb{D}$ of 3.2.7 by adding the axioms

$$\left(\top \vdash_{x_1,\ldots,x_n} (\exists y) \bigwedge_{i=1}^{n} (y \# x_i)\right)$$

for each natural number n (for $n = 0$, we interpret this as the sequent $(\top \vdash_{[]} (\exists y)\top)$). As we remarked after 3.2.8, we may construct a classifying topos for this theory by taking the classifying topos $[\mathbf{Set}_{fm}, \mathbf{Set}]$ for $\mathbb{D}$ and imposing a coverage J on $\mathbf{Set}_{fm}^{\mathrm{op}}$ 'generated' by the extra axioms of $\mathbb{D}_\infty$. Specifically, the nth axiom says that the inclusion map $n \rightarrowtail (n+1)$, where n as usual denotes the n-element set $\{0, 1, \ldots, n-1\}$, generates a J-covering cosieve on n; it is then easy to see that *every* morphism of $\mathbf{Set}_{fm}$ must generate a J-covering cosieve, and hence that J consists of all nonempty cosieves on objects of $\mathbf{Set}_{fm}$. So the classifying topos $\mathbf{Set}[\mathbb{D}_\infty] \simeq \mathbf{Sh}(\mathbf{Set}_{fm}^{\mathrm{op}}, J)$ is equivalent to $\mathbf{sh}_{\neg\neg}[\mathbf{Set}_{fm}, \mathbf{Set}]$; note that $(\mathbf{Set}_{fm}^{\mathrm{op}}, J)$ is an atomic site, as we saw in C3.5.9(c).

It is obvious that $\mathbb{D}_\infty$ is $\aleph_0$-categorical, since its models in $\mathbf{Set}$ are just infinite sets (and homomorphisms of models are injections between them). If we take its 'standard' countable model to be the set $\mathbb{N}$ of natural numbers, then the group which we obtain from it as in 3.4.8 is just the coherent group G discussed in 3.4.1. In fact it is not hard to give a direct proof that $\mathbf{Sh}(\mathbf{Set}_{fm}^{\mathrm{op}}, J)$ is equivalent to $\mathbf{Cont}(G)$: given a natural number n, let H_n denote the pointwise stabilizer of n (i.e. of $\{0, 1, \ldots, n-1\}$) in G, and let A_n be the set of left cosets of H_n with G acting by left translation. Clearly, A_n is a continuous G-set; and since every open subgroup of G contains a conjugate of H_n for some n, we see that the atoms of $\mathbf{Cont}(G)$ are precisely the quotients of the A_n. Given any injection $f : m \rightarrowtail n$, we may extend f to a permutation f' of $\mathbb{N}$, uniquely up to right multiplication by

elements of H_m, and hence we obtain a well-defined G-equivariant map $A_n \to A_m$ sending gH_n to $gf'H_m$; it is easy to see that every G-equivariant map $A_n \to A_m$ is of this form, and hence that the full subcategory of $\mathbf{Cont}(G)$ on the objects A_n is equivalent to $\mathbf{Set}_{fm}^{\mathrm{op}}$. Given that the A_n form a separating set for $\mathbf{Cont}(G)$, the claimed equivalence now follows from the Comparison Lemma (C2.2.3).

Example 3.4.11 Our second example is the theory $\mathbb{L}_\infty$ of dense linearly ordered objects without endpoints. This is defined over a single-sorted signature Σ having one binary relation-symbol $<$ apart from equality; we define the theory $\mathbb{L}$ of (decidably) linearly ordered objects to have axioms

$$(((x < y) \wedge (y < x)) \vdash_{x,y} \bot) \text{ and}$$

$$(\top \vdash_{x,y} ((x = y) \vee (x < y) \vee (y < x))) \ ,$$

and $\mathbb{L}_\infty$ is obtained from $\mathbb{L}$ by adding the further axioms $(\top \vdash_{[]} (\exists x)\top)$,

$$((x < y) \vdash_{x,y} (\exists z)((x < z) \wedge (z < y))) \text{ and}$$

$$(\top \vdash_x (\exists y, z)((y < x) \wedge (x < z))) \ .$$

(We could have added the axioms which say that $<$ is irreflexive and transitive, but they are both deducible from the axioms of $\mathbb{L}$.) We note that $\mathbb{L}$ is a disjunctive theory in the sense of 1.3.6; as we did for $\mathbb{D}$ in 3.2.7, we may identify its classifying topos with a functor category, namely $[\mathbf{Ord}_{fm}, \mathbf{Set}]$, where $\mathbf{Ord}_{fm}$ is the category of finite ordinals and order-preserving injections between them (equivalently, the category of finite models of $\mathbb{L}$ in $\mathbf{Set}$). As in the previous example, we may generate a coverage J on $\mathbf{Ord}_{fm}^{\mathrm{op}}$ from the additional axioms of $\mathbb{L}_\infty$: the first one says that the injection $0 \rightarrowtail 1$ generates a covering cosieve on 0, the second that the injection $2 \rightarrowtail 3$ which omits the middle element of 3 generates a covering cosieve on 2, and the third that the injection $1 \rightarrowtail 3$ which corresponds to the middle element also generates a covering cosieve. It is then straightforward to verify that each individual morphism of $\mathbf{Ord}_{fm}$ must generate a covering cosieve, so that J consists of all nonempty cosieves, and hence the classifying topos $\mathbf{Set}[\mathbb{L}_\infty] \simeq \mathbf{Sh}(\mathbf{Ord}_{fm}^{\mathrm{op}}, J)$ is equivalent to $\mathbf{sh}_{\neg\neg}[\mathbf{Ord}_{fm}, \mathbf{Set}]$.

As before, the theory $\mathbb{L}_\infty$ is well known to be $\aleph_0$-categorical: the fact that every countable dense linearly ordered set without endpoints is isomorphic to the ordered set $(\mathbb{Q}, <)$ of rational numbers is probably the best-known categoricity result in mathematics. So we may also identify $\mathbf{Set}[\mathbb{L}_\infty]$ with $\mathbf{Cont}(G)$ where G is the group of order-preserving permutations of $\mathbb{Q}$, topologized so that the (pointwise) stabilizers of finite subsets of $\mathbb{Q}$ form a base of neighbourhoods of the identity; the fact that G is a coherent topological group (i.e. that these subgroups have finite bi-index) may be proved by methods like those of 3.4.1. And we may also establish the equivalence $\mathbf{Sh}(\mathbf{Ord}_{fm}^{\mathrm{op}}, J) \simeq \mathbf{Cont}(G)$ directly, by identifying each object n of $\mathbf{Ord}_{fm}$ with the action of G on the left cosets of the stabilizer of an n-element set. Details of the proof can be found in [1085].

Note that if we had omitted the nonemptiness axiom $(\top \vdash (\exists x)\top)$ from our definition of $\mathbb{L}_\infty$, we should have obtained a theory which satisfies (b) and (a_1) but not (a_2): it has just two completions, both of which are $\aleph_0$-categorical (one is $\mathbb{L}_\infty$, and the other is the theory whose only model is empty), but the formula $\neg(\exists x)\top$ is not provably equivalent in it to any existential formula (and the inclusion $\emptyset \rightarrowtail \mathbb{Q}$ is an embedding of **Set**-models of the theory which is not elementary).

Example 3.4.12 For completeness, we also discuss the topos of sheaves on the atomic site $(\mathbf{Set}_{fe}, J)$ of finite sets and surjective functions, mentioned in C2.1.12(c) and C3.5.9(c), as a classifying topos. The Stone duality between $\mathbf{Set}_f$ and the category $\mathbf{Bool}_f$ of finite Boolean algebras restricts to an equivalence between $\mathbf{Set}_{fe}^{\mathrm{op}}$ and the category $\mathbf{Bool}_{fm}$ of finite Boolean algebras and injections. Now the functor category $[\mathbf{Bool}_f, \mathbf{Set}]$ is a classifying topos for the algebraic theory of Boolean algebras, by 3.1.2; and arguments like those of 3.2.7 show that $[\mathbf{Set}_{fe}^{\mathrm{op}}, \mathbf{Set}] \cong [\mathbf{Bool}_{fm}, \mathbf{Set}]$ classifies decidable Boolean algebras, that is Boolean algebras equipped with a binary relation # and axioms saying that # is the complement of the diagonal. And imposing the coverage on $\mathbf{Bool}_{fm}^{\mathrm{op}}$ in which every inhabited sieve covers corresponds to adding the axiom which says that the Boolean algebra is *atomless*, i.e.

$$((0 \# x) \vdash_x (\exists y)((y \leq x) \wedge (0 \# y) \wedge (y \# x))) .$$

Once again, it is well known that the theory of atomless Boolean algebras is complete and $\aleph_0$-categorical; its unique countable model is the free Boolean algebra on a countable infinity of generators, or equivalently the algebra of clopen subsets of the Cantor space K. Thus we may conclude that $\mathbf{Sh}(\mathbf{Set}_{fe}, J)$ is equivalent to the topos $\mathbf{Cont}(G)$, where G is the group of automorphisms of this Boolean algebra, topologized as usual so that the pointwise stabilizers of finite subalgebras form a base of neighbourhoods of the identity. Equivalently, by Stone duality, we may regard G as the group of self-homeomorphisms of the Cantor space, equipped with a topology which turns out to be generated by sets of the form

$$\{g \in G \mid g \in U \text{ and } g^{-1} \in V\}$$

where U and V are open sets in the usual compact-open topology on the set of all continuous maps $K \to K$. (The compact-open topology itself does not restrict to a group topology on G; the operation $g \mapsto g^{-1}$ is not continuous, which is why we have to modify the topology as above.)

Finally in this section, we consider how much of the theory developed above survives when we generalize from coherent to geometric theories. The answer, unfortunately, is 'not very much'. Lemma 3.4.5 remains true with 'coherent' replaced by 'geometric' throughout, and with a simpler proof – the reason is that every subobject in $\mathbf{Set}[\mathbb{T}] \simeq \mathcal{G}_{\mathbb{T}}$ of an object of $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ lies in $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$, and so the Booleanness of $\mathbf{Set}[\mathbb{T}]$ implies that $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ has Boolean subobject lattices.

Hence, by the easy direction of 3.4.9(i), we can deduce that *if* $\mathbb{T}$ *is a geometric theory whose classifying topos is Boolean, then every homomorphism of* $\mathbb{T}$*-models in* **Set** *is an embedding.* But the absence of the finiteness result 3.4.4 makes it much harder to prove any substantial results in the opposite direction.

From another point of view, the difficulty is that, once we have removed the restriction to coherent toposes, the properties of Booleanness and atomicity no longer coincide. In fact we can give a syntactic condition (albeit a less useful one than 3.4.6) for an arbitrary geometric theory $\mathbb{T}$ to have an atomic classifying topos: we now describe this. Given $\mathbb{T}$, we say that a geometric formula-in-context $\vec{x}\,.\,\phi$ is $\mathbb{T}$-*complete* (or simply *complete*) if the sequent $(\phi \vdash_{\vec{x}} \bot)$ is not provable in $\mathbb{T}$, but for every geometric formula ψ in the same context either $(\phi \vdash_{\vec{x}} \psi)$ or $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$ is provable. This is clearly equivalent to saying that the theory obtained from $\mathbb{T}$ by adding a string of new constants $\vec{c}$ of the same length and type as $\vec{x}$, plus the extra axiom $(\top \vdash_{[]} \phi[\vec{c}/\vec{x}])$, is a complete theory; in particular, $\mathbb{T}$ itself is complete iff $[\,]\,.\,\top$ is $\mathbb{T}$-complete.

By an *atomic system* for a geometric theory $\mathbb{T}$, we mean a function assigning to each context $\vec{x}$ over the signature of $\mathbb{T}$ a set $A_{\vec{x}}$ of geometric formulae in that context, such that

(i) $(\top \vdash_{\vec{x}} \bigvee_{\phi \in A_{\vec{x}}} \phi)$ is provable in $\mathbb{T}$, for each $\vec{x}$;

(ii) for any two distinct members ϕ and ψ of $A_{\vec{x}}$, the sequent $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$ is provable in $\mathbb{T}$;

(iii) if $\phi \in A_{\vec{x},y}$, then either ϕ is inconsistent or there is a member of $A_{\vec{x}}$ $\mathbb{T}$-provably equivalent to $(\exists y)\phi$;

(iv) if $\phi \in A_{\vec{x}}$ and ψ is any atomic formula in the context $\vec{x}$, then either $(\phi \vdash_{\vec{x}} \psi)$ or $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$ is provable in $\mathbb{T}$.

Proposition 3.4.13 *For a geometric theory* $\mathbb{T}$*, the following are equivalent:*

(i) *The classifying topos* **Set**$[\mathbb{T}]$ *is atomic (over* **Set***).*

(ii) *For every context* $\vec{x}$ *over the signature of* $\mathbb{T}$*, there is a set* $B_{\vec{x}}$ *of* $\mathbb{T}$*-complete formulae in that context such that* $(\top \vdash_{\vec{x}} \bigvee_{\phi \in B_{\vec{x}}} \phi)$ *is provable in* $\mathbb{T}$.

(iii) $\mathbb{T}$ *has an atomic system.*

Proof (i) $\Leftrightarrow$ (ii): By 1.4.4(iv), condition (ii) is equivalent to saying that, for every context $\vec{x}$, the subobject lattice of $\{\vec{x}\,.\,\top\}$ in $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ is atomic (and hence Boolean). This is clearly implied by (i); but it easily implies that every subobject lattice in $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ is atomic, and hence we may obtain the same result for the ∞-pretopos $\mathcal{G}_{\mathbb{T}}$, i.e. for **Set**$[\mathbb{T}]$. (Alternatively, we could apply the Comparison Lemma (C2.2.3) to the inclusion in $\mathcal{C}_{\mathbb{T}}^{\mathrm{geom}}$ of its full subcategory whose objects are $\mathbb{T}$-complete formulae-in-context, and then observe that the latter, with its induced coverage, is an atomic site as defined in C3.5.8.)

(ii) $\Leftrightarrow$ (iii): Clearly, if we are given a family of sets $B_{\vec{x}}$ as in (ii), we may obtain an atomic system from them by eliminating duplication from each $B_{\vec{x}}$: that is, we define $A_{\vec{x}}$ to contain one member from each provable-equivalence class of formulae in $B_{\vec{x}}$. Conversely, if we are given an atomic system of sets $A_{\vec{x}}$,

we obtain a family of sets $B_{\vec{x}}$ as in (ii) by eliminating any inconsistent formulae (that is, formulae ϕ for which $(\phi \vdash_{\vec{x}} \bot)$ is provable) from each $A_{\vec{x}}$: to prove that this works, we must show that each consistent member of $A_{\vec{x}}$ is $\mathbb{T}$-complete, i.e. that if $\phi \in A_{\vec{x}}$ and $\vec{x}.\psi$ is any geometric formula-in-context then either $(\phi \vdash_{\vec{x}} \psi)$ or $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$ is provable. We do this by induction on the structure of ψ: for atomic ψ, it is just clause (iv) of the definition of an atomic system. If $\psi = \bigvee_{i \in I} \theta_i$, then by induction we have either $(\phi \vdash_{\vec{x}} \theta_i)$ for some i (whence $(\phi \vdash_{\vec{x}} \psi)$) or $((\phi \wedge \theta_i) \vdash_{\vec{x}} \bot)$ for all i (whence $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$). Similarly if ψ is a finite conjunction. Finally, suppose $\psi = (\exists y)\theta$. Let $A^{\phi}_{\vec{x},y}$ be the set of formulae χ in $A_{\vec{x},y}$ for which $(\exists y)\chi$ is provably equivalent to ϕ; we note that if χ' is any other member of $A_{\vec{x},y}$, then $(\exists y)\chi'$ is inconsistent with ϕ by conditions (ii) and (iii) of the definition, from which we may deduce $((\phi \wedge \chi') \vdash_{\vec{x},y} \bot)$. Hence by condition (i) ϕ is provably equivalent (in the context $\vec{x}, y$) to the disjunction of all the members of $A^{\phi}_{\vec{x},y}$. It now follows easily that if some member χ of $A^{\phi}_{\vec{x},y}$ satisfies $(\chi \vdash_{\vec{x},y} \theta)$ then we have $(\phi \vdash_{\vec{x}} \psi)$, and if none of them do so we have $((\phi \wedge \theta) \vdash_{\vec{x},y} \bot)$, whence $((\phi \wedge \psi) \vdash_{\vec{x}} \bot)$. □

We note that, in the context of 3.4.13, $\mathbb{T}$ is (equivalent to) a coherent theory iff the sets $A_{\vec{x}}$ of 3.4.13(iii) may be taken to be finite for all $\vec{x}$. Also, $\mathbb{T}$ is complete iff (it is consistent and) $A_{[]}$ may be taken to be the singleton $\{\top\}$.

Example 3.4.14 As an application of 3.4.13, we construct an example of a connected atomic topos with no points, as promised in C5.2.14(a). The topos will be constructed as **Set**$[\mathbb{T}]$, where $\mathbb{T}$ is a complete theory satisfying the conditions of 3.4.13 but having no models in **Set**: the theory in question was first studied by J. Malitz [793] (see also [777]). We take a signature Σ having two sorts A and B, a constant $a_0 \colon 1 \to A$, three binary relation symbols $\# \rightarrowtail A, A$, $\prec \rightarrowtail A, A$ and $\in \rightarrowtail A, B$ (as usual, we write these in infix notation), and a family of unary relation symbols $R_i \rightarrowtail B$, $i \in I$, where I is a set whose cardinality is strictly greater than that of the continuum. For convenience, we shall also use the notation $\overline{\prec}$ for the transitive closure of $\prec$, i.e.

$$\Big((x \mathrel{\overline{\prec}} y) \dashv\vdash \bigvee_{n \in \mathbb{N}} (\exists z_1, \ldots, z_n)((x \prec z_1) \wedge (z_1 \prec z_2) \wedge \cdots \wedge (z_n \prec y))\Big)$$

(the case $n = 0$ of this disjunction being interpreted as $(x \prec y)$).

The first group of axioms states that, in any model M, the interpretations of $\#$ and $\prec$ make MA into a decidable complete binary tree with root Ma_0:

$$\begin{gathered}
(\top \vdash_{x,y} ((x = y) \vee (x \# y))) \\
((x \# x) \vdash_x \bot) \\
(\top \vdash_x ((a_0 = x) \vee (a_0 \mathrel{\overline{\prec}} x))) \\
(((y \prec x) \wedge (z \prec x)) \vdash_{x,y,z} (y = z)) \\
(\top \vdash_x (\exists y, z)((x \prec y) \wedge (x \prec z) \wedge (y \# z))) \\
(((x \prec y) \wedge (x \prec z) \wedge (x \prec w)) \vdash_{x,y,z,w} ((y = z) \vee (z = w) \vee (w = y))) \\
((x \mathrel{\overline{\prec}} y) \vdash_{x,y} (x \# y)) .
\end{gathered}$$

The next group says that, for each element b of MB, the set $\{a \in MA \mid (a,b) \in [\![\in]\!]_M\}$ is a (full) branch of the tree MA, and that distinct elements of MB yield distinct branches:

$$\begin{gathered}(\top \vdash_u (a_0 \in u))\\ (((x \in u) \wedge (y \prec x)) \vdash_{u,x,y} (y \in u))\\ ((x \in u) \vdash_{x,u} (\exists y)((x \prec y) \wedge (y \in u)))\\ (((x \in u) \wedge y \in u)) \vdash_{u,x,y} ((x = y) \vee (x \,\overline{\prec}\, y) \vee (y \,\overline{\prec}\, x)))\\ (\top \vdash_{u,v} ((u = v) \vee (\exists x,y,z)((x \prec y) \wedge (x \prec z) \wedge (y \# z) \wedge (y \in u) \wedge (z \in v)))) .\end{gathered}$$

Here the variables u and v are of sort B, and x, y, z are of sort A. Finally, we write down axioms saying that MB is partitioned into the sets MR_i, and that each of these is dense in the set of all branches of MA (topologized as the Cantor set):

$$\begin{array}{ll}(\top \vdash_u \bigvee_{i \in I} R_i(u)) & \\ ((R_i(u) \wedge R_j(u)) \vdash_u \bot) & (i \neq j)\\ (\top \vdash_x (\exists u)((x \in u) \wedge R_i(u))) & (\text{for each } i) .\end{array}$$

It is clear that $\mathbb{T}$ cannot have any models in **Set**; for in any **Set**-model of the first two groups of axioms MB would have cardinality at most that of the continuum, and hence could not be partitioned into an I-indexed family of nonempty subsets. On the other hand, if I were a countably infinite set, then we could construct a model of $\mathbb{T}$; and since the geometric propositional theory of bijections $I \to \mathbb{N}$ is consistent (cf. the remarks after C1.2.9), we can construct a non-degenerate Boolean **Set**-topos in which $\mathbb{T}$ has a model; so in particular it is consistent.

To verify that $\mathbb{T}$ has an atomic system in the sense of 3.4.13, we shall introduce a number of derived predicates. Let $C_n(x)$ $(n \in \mathbb{N})$ denote the predicate asserting that x sits at 'level n' in the binary tree, i.e. (for $n \geq 2$) the formula $(\exists y_1, \ldots, y_{n-1})((a_0 \prec y_1) \wedge (y_1 \prec y_2) \wedge \cdots \wedge (y_{n-1} \prec x))$. Similarly, let $D_n(x,y)$ denote the assertion that x and y have a common predecessor at level n in the tree, but not at level $n+1$; let $E_n(x,u)$ denote the assertion that the level-n member of the branch u is a predecessor of x (or equal to x) but its level-$(n+1)$ member is not, and let $F_n(u,v)$ be the assertion that the branches u and v have a common member at level n but not at level $n+1$. It is easily verified that, from the axioms given above, we can deduce things like

$$\Big(\top \vdash_{u,v} ((u = v) \vee \bigvee_{n \in \mathbb{N}} F_n(u,v))\Big)$$

and that the disjunction on the right-hand side of this sequent is provably disjoint. Now, given a context $x_1, x_2, \ldots, x_r, u_1, u_2, \ldots, u_s$ over Σ (with the x_j of sort A and the u_k of sort B), we define a *pattern* in this context to be a finite conjunction of formulae which includes

(i) for each $j \leq r$, just one of the formulae $C_n(x_j)$;

(ii) for each pair (j_1, j_2), just one of the formulae $D_n(x_{j_1}, x_{j_2})$;

(iii) for each $k \leq s$, just one of the formulae $R_i(u_k)$;
(iv) for each pair (k_1, k_2), either the formula $(u_{k_1} = u_{k_2})$ or one of the formulae $F_n(u_{k_1}, u_{k_2})$; and
(v) for each pair (j, k), just one of the formulae $E_n(x_j, u_k)$.

Clearly, in any context with more than one variable there will be many possible patterns which are inconsistent; but since we allowed inconsistent formulae to appear in the members of an atomic system, this does not matter. We define $A_{\vec{x},\vec{u}}$ to be the set of all possible patterns in the context $\vec{x}, \vec{u}$: the verification that this is indeed an atomic system is long but entirely straightforward. We note also that the only possible pattern in the empty context is the empty conjunction $\top$; hence **Set**$[\mathbb{T}]$ is not only atomic but connected.

Suggestions for further reading: Blass & Ščedrov [133], Bunge & Reyes [206], Makkai [777], Ščedrov [1085].

D3.5 Conceptual completeness

The classical completeness theorem says that if something is true in all models of a theory, then it should be provable in the theory. But there is a stonger notion of completeness, which asserts that if something exists (in a 'natural' way) in all models of a theory, then 'the theory should have a name for it'. The functional completeness result 3.1.3 is an example of this type of completeness. More generally, we should like to prove that if every model M comes equipped with some particular subobject $R_M \rightarrowtail M(A_1, \ldots, A_n)$ (where $A_1, \ldots, A_n$ are sorts of the appropriate signature), and every homomorphism of models preserves this relation, then there should exist some formula-in-context $\vec{x} \,.\, \phi$ (where $x_i \colon A_i$ for each i) whose interpretation in each model M is precisely R_M. (To take a simple example: the notion of semilattice may be defined – as a commutative idempotent monoid – in a way which makes no mention of a partial ordering; but from the fact that every semilattice possesses a canonical partial order, and every semilattice homomorphism is order-preserving, we should expect the ordering to be definable by a formula-in-context over the signature for monoids – and indeed it is, by the formula $x, y \,.\, (mxy = x)$, where m is the monoid operation.)

If we interpret 'every model M' as meaning 'every model in a category of the appropriate kind', then this result, like the ordinary completeness result, becomes an easy consequence of the existence of the syntactic category; for we saw in 1.4.4(iv) that every subobject of an object of the form $M_{\mathbb{T}}(A_1, \ldots, A_n) = \{\vec{x} \,.\, \top\}$ is the interpretation in $M_{\mathbb{T}}$ of some formula-in-context $\vec{x} \,.\, \phi$. But, if we wish to deduce definability results merely from the existence of relations on $\mathbb{T}$-models in **Set**, then we have some further work to do. The following proof may remind the reader of the proof which we gave for 3.4.9(i); indeed, the latter may be seen as a special case of it.

Theorem 3.5.1 (Definability Theorem for coherent logic) *Let $\mathbb{T}$ be a coherent theory over a signature Σ, let $A_1 \cdots A_n$ be a string of sorts of Σ, and suppose we are given, for every* **Set***-model M of $\mathbb{T}$, a distinguished subset R_M of $M(A_1, \ldots, A_n)$ in such a way that each $\mathbb{T}$-model homomorphism $h \colon M \to N$ in* **Set** *maps R_M into R_N. Then there exists a coherent formula-in-context $\vec{x}.\phi$ over Σ, where $x_i \colon A_i$ for each i, such that $[\![\vec{x}.\phi]\!]_M = R_M$ for each M.*

Proof First we extend Σ to a new signature Σ_1 by adding a new relation symbol $R \rightarrowtail A_1 \cdots A_n$; then we may make each $\mathbb{T}$-model M in **Set** into a Σ_1-structure by interpreting the new symbol as R_M, and we define $\mathbb{T}_1$ to be the set of all coherent sequents over Σ_1 which are satisfied in all such structures. Clearly, $\mathbb{T}_1$ includes all the axioms of $\mathbb{T}$. Next, we define Γ to be the set of all coherent formulae γ over Σ, in the particular context $\vec{x}$, for which the sequent $(\gamma \vdash_{\vec{x}} R(\vec{x}))$ is in $\mathbb{T}_1$. We need to show that there is some $\gamma \in \Gamma$ for which $(R(\vec{x}) \vdash_{\vec{x}} \gamma)$ is also in $\mathbb{T}_1$.

Suppose not; then for each γ we can find a $\mathbb{T}_1$-model M and a string of elements $m_i \in MA_i$ such that $(m_1, \ldots, m_n)$ belongs to R_M but not to the interpretation of $\vec{x}.\gamma$ in M. Now let us enlarge Σ_1 to Σ_2 by adding a string of new constants $c_1, \ldots, c_n$ of sorts $A_1, \ldots, A_n$, and let $\mathbb{T}_2$ be the theory over Σ_2 consisting of the axioms of $\mathbb{T}_1$ plus all sequents $(\gamma[\vec{c}/\vec{x}] \vdash_{[]} \bot)$ with $\gamma \in \Gamma$, and the single sequent $(\top \vdash_{[]} R(\vec{c}))$. We claim that $\mathbb{T}_2$ is consistent; for any derivation of $(\top \vdash_{[]} \bot)$ from it would involve only finitely many of the axioms $(\gamma[\vec{c}/\vec{x}] \vdash_{[]} \bot)$, and since finite disjunctions of coherent formulae are coherent, the first sentence of this paragraph assures us that we can find a model for this subtheory of $\mathbb{T}_2$.

Let M_0 be a model of $\mathbb{T}_2$. Now let Σ_3 be the signature obtained from Σ_2 by adding a new constant $d_m \colon A$ for each $m \in M_0A$, as A ranges over all the sorts of Σ; we make M_0 into a Σ_3-structure in the obvious way. Let Δ be the set of all atomic sentences over Σ_3 which are true in M_0, and consider the theory $\mathbb{T}_3$ over Σ_3 which consists of the axioms of $\mathbb{T}_1$, all sequents $(\top \vdash_{[]} \delta)$ with $\delta \in \Delta$, and the sequent $(R(\vec{c}) \vdash_{[]} \bot)$.

It is easy to see that a model of $\mathbb{T}_3$ is essentially the same as a model N of $\mathbb{T}_1$, together with a homomorphism $h \colon M_0 \to N$ (which sends an element $m \in M_0A$ to the interpretation of d_m in N), such that the n-tuple $(h(m_1), \ldots, h(m_n))$ (where m_i denotes the interpretation of c_i in M_0) does not lie in R_N. But since $(m_1, \ldots, m_n)$ *does* lie in R_M, the hypothesis of the theorem says that no such N can exist. So the theory $\mathbb{T}_3$ is inconsistent; that is, $(\top \vdash_{[]} \bot)$ is derivable from it. As before, it must in fact be derivable from a finite subset of $\mathbb{T}_3$; let $\delta_1, \ldots, \delta_k$ be the members of Δ for which $(\top \vdash \delta_i)$ is in this finite set, let δ denote their conjunction, and let $\overline{\delta}$ be the formula obtained from δ on replacing each constant d_m by a variable y_m of the appropriate sort (but leaving the constants c_i, if they occur). Since the d_m do not occur in $R(\vec{c})$ or in any of the axioms of $\mathbb{T}_1$, it is easy to see that we can in fact derive a contradiction from

$$\mathbb{T}_1 \cup \{(R(\vec{c} \vdash \bot), (\top \vdash (\exists \vec{y})\overline{\delta})\} .$$

This means that the sequent $((\exists \vec{y})\phi \vdash_{\vec{x}} R(\vec{x}))$, where ϕ is obtained from $\overline{\delta}$ on replacing the constants c_i by variables x_i, is satisfied in all $\mathbb{T}_1$-models; thus $\vec{x}\,.\,(\exists \vec{y})\phi$ belongs to the set Γ defined earlier. But, by construction, the n-tuple $(m_1, \ldots, m_n)$ belongs to the interpretation of this formula in M_0, whereas it does not belong to the interpretation of any member of Γ. This contradiction completes the proof of the theorem. □

Remarks 3.5.2 (*a*) The hypothesis, in the statement of 3.5.1, that we are given a choice of R_M for *every* $\mathbb{T}$-model M in **Set** can be relaxed: it is sufficient to have such a choice for every model in some given class $\mathcal{M}$ of models, and to know that these choices are compatible with all homomorphisms between models in $\mathcal{M}$ – provided $\mathcal{M}$ has suitable closure properties to ensure that the model N which we construct in the course of the proof of 3.5.1 belongs to it. (For example, it would be sufficient to suppose that $\mathcal{M}$ is a cosieve in $\mathbb{T}$-Mod(**Set**).). We shall then arrive at a coherent formula-in-context $\vec{x}\,.\,\phi$ whose interpretation in every model in $\mathcal{M}$ is our given subset, although we have no control over its interpretation in other models of $\mathbb{T}$ – and, in general, it will not be unique up to $\mathbb{T}$-provable equivalence (whereas, by classical completeness, it will be unique if we specify its interpretation in every $\mathbb{T}$-model).

(*b*) The definability theorem is not true for regular or cartesian logic, for an obvious reason: if we take R_M (for each model M) to be the union of two subsets defined by regular formulae-in-context, then it will be preserved by arbitrary homomorphisms, but will not in general be definable by a single regular formula (cf. 3.3.11). Nevertheless, we shall see later in this section that the conceptual completeness theorem, which for coherent logic may be viewed as a generalization of the result just proved, does hold for these two fragments of first-order logic. We should also mention that there is a 'classical definability theorem' for full (finitary) first-order logic, generally associated with the name of E. W. Beth, in which the hypothesis on the chosen subsets R_M is simply that they are preserved by arbitrary isomorphisms of $\mathbb{T}$-models in **Set**; see [446, 6.6.4 and 10.5.1].

One possible way of regarding Theorem 3.5.1 is as follows. Using the notation of the proof, we have an *interpretation* of $\mathbb{T}$ in $\mathbb{T}_1$, that is a functor I (preserving the appropriate structure) from the syntactic category of $\mathbb{T}$ to that of $\mathbb{T}_1$. Composition with I yields a functor $I^* : \mathbb{T}_1\text{-Mod}(\mathbf{Set}) \to \mathbb{T}\text{-Mod}(\mathbf{Set})$: of course, this is just the functor which 'forgets' the interpretation of the extra relation symbol R. This functor is part of an equivalence of categories; observe that each $\mathbb{T}$-model has a *unique* extension to a $\mathbb{T}_1$-model, since the 'naturality' hypothesis ensures that there is a unique choice for R_M in every M (the identity morphism on M is a $\mathbb{T}$-model homomorphism!). And the conclusion of the theorem says (roughly) that I itself is an equivalence of categories (or at least a weak equivalence, as we defined it in Section A1.1), since the 'new' object $\{\vec{x}\,.\,R(\vec{x})\}$ of $\mathcal{C}_{\mathbb{T}_1}$ is isomorphic to one in the image of I. (Incidentally, since everything we shall do in this section depends on the classical completeness theorem, and therefore requires the axiom of choice – or at least the Boolean Prime Ideal Theorem, cf. 1.5.11 – we

shall not worry too much about the difference between 'equivalence' and 'weak equivalence'.)

It would thus be of interest to have a theorem of the following form: 'Let $\mathcal{C}$ and $\mathcal{D}$ be small categories with structure (cartesian, regular or coherent, according to taste), and $I: \mathcal{C} \to \mathcal{D}$ a structure-preserving functor. If the functor $I^*: \mathfrak{K}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{K}(\mathcal{C}, \mathbf{Set})$ induced by composition with I is a weak equivalence (where $\mathfrak{K}$ denotes the meta-2-category of cartesian/regular/coherent categories), then I is a weak equivalence'. Such a result is known as a *conceptual completeness theorem* for cartesian/regular/coherent logic.

For cartesian logic, the theorem is indeed true in the form just stated, and even in a stronger form:

Lemma 3.5.3

(i) *Let $\mathcal{C}$ and $\mathcal{D}$ be small cartesian categories such that $\mathfrak{Cart}(\mathcal{C}, \mathbf{Set}) \simeq \mathfrak{Cart}(\mathcal{D}, \mathbf{Set})$. Then $\mathcal{C}$ and $\mathcal{D}$ are equivalent.*

(ii) *Let $\mathbb{S}$ and $\mathbb{T}$ be cartesian theories such that* $\mathbb{S}$-Mod($\mathbf{Set}$) $\simeq$ $\mathbb{T}$-Mod($\mathbf{Set}$). *Then $\mathbb{S}$ and $\mathbb{T}$ are Morita-equivalent.*

Proof (i) We recall from C4.2.2(v) and B3.2.5 that, for a small cartesian category $\mathcal{C}$, the category $\mathfrak{Cart}(\mathcal{C}, \mathbf{Set})$ may be identified with the free filtered-colimit completion Ind-$\mathcal{C}^{\mathrm{op}}$ of $\mathcal{C}^{\mathrm{op}}$; moreover, $\mathcal{C}^{\mathrm{op}}$ is fully embedded in this category by the Yoneda embedding, and its objects may be recovered as the finitely-presentable objects of Ind-$\mathcal{C}^{\mathrm{op}}$ by C4.2.2(ii) (note that, since $\mathcal{C}$ is cartesian, it is Cauchy-complete). So any equivalence Ind-$\mathcal{C}^{\mathrm{op}} \simeq$ Ind-$\mathcal{D}^{\mathrm{op}}$ must restrict to an equivalence $\mathcal{C} \simeq \mathcal{D}$.

(ii) The second statement of the lemma is a simple translation of the first into the language of first-order theories, by 1.4.7 and 1.4.8. □

Remark 3.5.4 For coherent logic, the conceptual completeness theorem cannot be true in this simple-minded form: there can be equivalences $\mathbb{S}$-Mod($\mathbf{Set}$) $\simeq$ $\mathbb{T}$-Mod($\mathbf{Set}$) which are not induced by any functor between the syntactic categories of $\mathbb{S}$ and $\mathbb{T}$. For a counterexample, consider the propositional theory $\mathbb{T}_B$ of prime filters of a Boolean algebra B, defined as in 1.1.7(m). Since the prime filters of a Boolean algebra are all maximal (cf. 4.6.16 below), the category of $\mathbf{Set}$-models of $\mathbb{T}_B$ is (equivalent to) the discrete category whose objects are the prime filters. So an equivalence between two such categories, for Boolean algebras B_1 and B_2, would simply say that their spectra have the same cardinality; whereas, by 3.3.14, Morita equivalence of the theories would say that the spectra of B_1 and B_2 are homeomorphic (and hence, by Stone duality, that B_1 and B_2 are isomorphic). It is easy to give examples of pairs of Boolean algebras satisfying the first condition but not the second.

Indeed, for regular and coherent logic, the conceptual completeness theorem does not even hold in the form given before 3.5.3. For if $\mathcal{C}$ is any non-effective small regular category, then the embedding of $\mathcal{C}$ in its effectivization is not a weak

equivalence, but it induces an equivalence $\mathfrak{Reg}(\mathcal{C}, \mathbf{Set}) \simeq \mathfrak{Reg}(\mathbf{Eff}(\mathcal{C}), \mathbf{Set})$ by A3.3.10, since **Set** is effective regular. Similarly for the embedding of any small coherent category $\mathcal{C}$ in its pretopos completion $\mathbf{Eff}(\mathbf{Pos}(\mathcal{C}))$. Thus the best we can hope for is a conceptual completeness theorem in which we restrict the categories $\mathcal{C}$ and $\mathcal{D}$ to be effective regular in the first case, and pretoposes in the second. (This is of course an important part of the reason why we defined Morita equivalence for regular and coherent theories in the way in which we did.) However, we note that this does not affect the definability result from which we started in this section: for the processes of effectivizing a regular category, and of positivizing a coherent category, do not add any new subobjects of the objects of the original category. (In the terminology to be introduced below, the embedding of a regular category in its effectivization, or of a coherent category in its positivization, is full on subobjects.)

In what follows, we shall (as in Section D3.3) deal almost exclusively with the coherent case; but the theorem is also true in the regular case, and may be proved by very similar methods. We shall discuss the changes which need to be made for the regular case in 3.5.12 below. In fact we may break the theorem down into a series of lemmas, each of which derives properties of the functor I from 'dual' properties of I^*. For the first of these, the assumption that we are dealing with pretoposes rather than arbitrary coherent categories is not needed:

Lemma 3.5.5 *Let $\mathcal{C}$ and $\mathcal{D}$ be small coherent categories, and $I: \mathcal{C} \to \mathcal{D}$ a coherent functor. If the functor $I^*: \mathfrak{Coh}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{Coh}(\mathcal{C}, \mathbf{Set})$ induced by composition with I is essentially surjective on objects, then I is conservative.*

Proof Let $\mathbb{T}_{\mathcal{C}}$ and $\mathbb{T}_{\mathcal{D}}$ denote the (coherent) theories of coherent functors defined on $\mathcal{C}$ and $\mathcal{D}$, as constructed after 1.4.11. Let $m: A' \rightarrowtail A$ be a monomorphism in $\mathcal{C}$ such that $I(m)$ is an isomorphism. Then the sequent

$$(\top \vdash_x (\exists y)(m(y) = x))$$

(where x and y are variables of sorts A and A' respectively) is satisfied in any $\mathbb{T}_{\mathcal{C}}$-model which is the image under I^* of a $\mathbb{T}_{\mathcal{D}}$-model. But the hypothesis of the lemma says that *every* $\mathbb{T}_{\mathcal{C}}$-model in **Set** is isomorphic to such a model; so by the completeness theorem 1.5.10(ii) it is provable in $\mathbb{T}_{\mathcal{C}}$, and hence satisfied in the model which corresponds to the identity functor on $\mathcal{C}$. In other words, m is a cover, and hence an isomorphism, in $\mathcal{C}$. So it follows from A1.2.4 that I is conservative. □

The next lemma is really the heart of the proof of conceptual completeness (as may be seen from the fact that the definability theorem is used in proving it). To state it, we need a new definition: we say that a functor $I: \mathcal{C} \to \mathcal{D}$ (which is assumed to preserve monomorphisms) is *full on subobjects* if, given any object A of $\mathcal{C}$ and a subobject $B \rightarrowtail IA$ in $\mathcal{D}$, there exists $A' \rightarrowtail A$ in $\mathcal{C}$ such that $IA' \cong B$ in $\mathrm{Sub}_{\mathcal{D}}(IA)$.

Lemma 3.5.6 *Let $\mathcal{C}$ and $\mathcal{D}$ be small coherent categories, and $I: \mathcal{C} \to \mathcal{D}$ a coherent functor. If the induced functor $I^*: \mathfrak{Coh}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{Coh}(\mathcal{C}, \mathbf{Set})$ is full, then I is full on subobjects. If in addition I^* is essentially surjective on objects, then I is full.*

Proof Suppose given a monomorphism $B \rightarrowtail IA$ in $\mathcal{D}$. As before, we let $\mathbb{T}_{\mathcal{C}}$ and $\mathbb{T}_{\mathcal{D}}$ denote the theories of coherent functors defined on $\mathcal{C}$ and $\mathcal{D}$. If I^* is not essentially surjective on objects, then not every $\mathbb{T}_{\mathcal{C}}$-model is of the form $I^*(N)$, where N is a $\mathbb{T}_{\mathcal{D}}$-model; but, as we remarked in 3.5.2(a), this does not matter. For any model M which is of this form, we may choose a $\mathbb{T}_{\mathcal{D}}$-model N with $I^*(N) \cong M$ and then define $R_M \subseteq MA$ to be (the image of) the composite $NB \rightarrowtail NIA \cong MA$; then the hypothesis of the lemma ensures that this choice of subobjects is preserved by arbitrary $\mathbb{T}_{\mathcal{C}}$-model homomorphisms. So by the definability theorem 3.5.1 there is a coherent formula-in-context $x\,.\,\phi$ (where $x : A$) such that, if $A' \rightarrowtail A$ denotes the interpretation of this formula in the $\mathbb{T}_{\mathcal{C}}$-model corresponding to the identity functor on $\mathcal{C}$, then $IA' \rightarrowtail IA$ and $B \rightarrowtail IA$ are mapped to isomorphic subobjects by every coherent functor $\mathcal{D} \to \mathbf{Set}$. An easy application of classical completeness then shows that we must have $IA' \cong B$ in $\mathrm{Sub}(IA)$.

For the second assertion, we have simply to note that if a cartesian functor I is both conservative and full on subobjects then it must be full. For if $g: IA \to IB$ is any morphism in $\mathcal{D}$, its graph is a subobject of $I(A \times B)$, and so must be isomorphic to IA' for some $(d, e): A' \rightarrowtail A \times B$ in $\mathcal{C}$. But then the conservativity of I implies that d is an isomorphism, so this subobject is the graph of a morphism $f = ed^{-1}$ in $\mathcal{C}$, which clearly satisfies $If = g$. □

For the next stage of the proof, we again need a new definition. Let us say that a coherent functor $I: \mathcal{C} \to \mathcal{D}$ is *finitely covering* if, for each object B of $\mathcal{D}$, we can find a finite family of objects $(A_1, \ldots, A_n)$ of $\mathcal{C}$, subobjects $C_j \rightarrowtail IA_j$ in $\mathcal{D}$, and morphisms $C_j \to B$ in $\mathcal{D}$ the union of whose images is the whole of B.

Lemma 3.5.7 *Let $\mathcal{C}$ and $\mathcal{D}$ be small coherent categories, and $I: \mathcal{C} \to \mathcal{D}$ a coherent functor such that $I^*: \mathfrak{Coh}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{Coh}(\mathcal{C}, \mathbf{Set})$ is faithful. Then I is finitely covering.*

Proof Let B be an object of $\mathcal{D}$, and consider the set X of pairs (A, ϕ) where A is an object of $\mathcal{C}$ and ϕ is a coherent formula over the signature of $\mathbb{T}_{\mathcal{D}}$, having free variables x and y of sorts IA and B respectively, for which it is provable in $\mathbb{T}_{\mathcal{D}}$ that 'ϕ denotes a partial map from IA to B'; that is, the sequent

$$((\phi \wedge \phi[y'/y]) \vdash_{x,y,y'} (y = y'))$$

is provable in $\mathbb{T}_{\mathcal{D}}$. Add a new constant c of sort B to the signature of $\mathbb{T}_{\mathcal{D}}$, and consider the theory $\mathbb{T}_1$ over this enlarged signature whose axioms are those of $\mathbb{T}_{\mathcal{D}}$ plus all axioms of the form $(\phi[c/y] \vdash_x \bot)$ where $(A, \phi) \in X$. If I is *not* finitely covering at the object B, then it is easy to see that each finite subset of $\mathbb{T}_1$ has a model, and hence $\mathbb{T}_1$ is consistent.

Consider a particular model of $\mathbb{T}_1$; we regard it as a pair (M, m_0) where M is a coherent functor $\mathcal{D} \to \mathbf{Set}$ and m_0 is the particular element of MB which interprets the constant c. Our aim now is to construct another coherent functor $N\colon \mathcal{D} \to \mathbf{Set}$ and two natural transformations $\alpha, \beta\colon M \rightrightarrows N$ which agree at all objects of the form IA but satisfy $\alpha_B(m_0) \neq \beta_B(m_0)$; this will contradict the hypothesis that I^* is faithful, and so complete the proof. Much as in the proof of 3.5.1, we construct the triple (N, α, β) as a model of a suitable coherent theory which we can show to be consistent.

First we enlarge our signature again, by adding two new constants d_m and e_m, both of sort C, for each element $m \in MC$ (and each object C of $\mathcal{D}$). Let Δ be the set of all sequents $(\top \vdash_{[]} \delta)$, where δ is an atomic sentence involving the d_m but not the e_m, which are satisfied in M when each d_m is interpreted as m, and let E be the set of similar sequents involving the e_m but not the d_m. Let Γ be the set of all sequents $(\top \vdash_{[]} (d_m = e_m))$ where the sort of d_m and e_m is in the image of I. Then it is clear that a model of the theory

$$\mathbb{T}_2 = \mathbb{T}_{\mathcal{D}} \cup \Gamma \cup \Delta \cup \mathrm{E}$$

is essentially the same thing as a coherent functor $N\colon \mathcal{D} \to \mathbf{Set}$ equipped with two natural transformations $M \rightrightarrows N$ (sending m to the interpretations of d_m and e_m respectively) which agree on objects of the form IA. It thus suffices to show that we may consistently add the sequent $((d_{m_0} = e_{m_0}) \vdash_{[]} \bot)$ to this theory.

Suppose some finite subset $\mathbb{T}_3$ of $\mathbb{T}_2 \cup \{((d_{m_0} = e_{m_0}) \vdash_{[]} \bot)\}$ is inconsistent. By adding more axioms to $\mathbb{T}_3$ if necessary, we may assume that it contains exactly the corresponding members of Δ and of E; also that it contains the sequent $(\top \vdash_{[]} (d_m = e_m))$ whenever d_m is mentioned in one of the members of Δ which appear in $\mathbb{T}_3$ and has sort in the image of I. Let δ be the conjunction of the atomic sentences δ_i such that $(\top \vdash_{[]} \delta_i)$ appears in $\mathbb{T}_3$; then we may replace all these sequents by a single sequent $(\top \vdash_{[]} \delta)$, and similarly we may replace the sequents in E which appear by a single sequent $(\top \vdash_{[]} \epsilon)$. Moreover, we may write $\delta = \phi[\vec{d_1}, \vec{d_2}, d_{m_0}/\vec{x}, \vec{y}, z]$ and $\epsilon = \phi[\vec{e}_1, \vec{e}_2, e_{m_0}/\vec{x}, \vec{y}, z]$, where ϕ is a formula in which we have substituted free variables for the constants appearing in δ, and we have separated the latter into three groups: those whose sorts are in the image of I, those other than d_{m_0} whose sorts are not in the image of I, and d_{m_0} itself. But now the failure of $\mathbb{T}_3$ to have a model tells us that every $\mathbb{T}_{\mathcal{D}}$-model must satisfy the sequent

$$(((\exists \vec{y})\phi \wedge (\exists \vec{y})\phi[z'/z]) \vdash_{\vec{x},z,z'} (z = z'))\ ;$$

in other words, it is provable in $\mathbb{T}_{\mathcal{D}}$ that the formula $(\exists \vec{y})\phi$ denotes a partial map $IA \rightharpoonup B$, where A is the product (in $\mathcal{C}$) of the sorts of the variables in $\vec{x}$. So the pair $(A, (\exists \vec{y})\phi)$ belongs to the set X that we constructed at the beginning of the proof – which contradicts the fact that we chose (M, m_0) to be a model of $\mathbb{T}_1$.

□

To complete the proof of the conceptual completeness theorem, we need

Lemma 3.5.8 *Let $\mathcal{C}$ be a pretopos, $\mathcal{D}$ a coherent category, and $I: \mathcal{C} \to \mathcal{D}$ a coherent functor which is full on subobjects, faithful and finitely covering. Then I is essentially surjective on objects.*

Proof Given any object B of $\mathcal{D}$, we have a finite covering family of partial maps $(IA_j \rightharpoonup B \mid 1 \leq j \leq n)$. Since I is full on subobjects, the domain of each member of this family may be taken to be of the form IA'_j for some $A'_j \rightarrowtail A_j$ in $\mathcal{C}$; then we may form the disjoint coproduct $A = \coprod_{j=1}^{n} A'_j$ in $\mathcal{C}$, and since I preserves disjoint coproducts we have a cover $IA \twoheadrightarrow B$ in $\mathcal{D}$. Let $S \rightarrowtail IA \times IA \cong I(A \times A)$ be the kernel-pair of this cover; then we may again write $S \cong IR$ for some $R \rightarrowtail A \times A$ in $\mathcal{C}$, and since I is faithful and hence conservative by A1.4.9, it is easy to see that R 'inherits' the property of being an equivalence relation from S. Let $A \twoheadrightarrow \overline{A}$ be a cover in $\mathcal{C}$ whose kernel-pair is $R \rightrightarrows A$; then $IA \twoheadrightarrow I\overline{A}$ and $IA \twoheadrightarrow B$ are both coequalizers of $S \rightrightarrows IA$ in $\mathcal{D}$, and hence $I\overline{A} \cong B$. □

Putting everything together, we now have

Theorem 3.5.9 (Conceptual Completeness for coherent logic)

(i) *Let $\mathcal{C}$ and $\mathcal{D}$ be small pretoposes, and $I: \mathcal{C} \to \mathcal{D}$ a coherent functor such that $I^*: \mathfrak{Coh}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{Coh}(\mathcal{C}, \mathbf{Set})$ is a weak equivalence. Then I is a weak equivalence.*

(ii) *Let $\mathbb{S}$ and $\mathbb{T}$ be coherent theories, and suppose we are given an interpretation of $\mathbb{S}$ in $\mathbb{T}$ for which the induced functor $\mathbb{T}$-Mod($\mathbf{Set}$) $\to$ $\mathbb{S}$-Mod($\mathbf{Set}$) is a weak equivalence. Then $\mathbb{S}$ and $\mathbb{T}$ are Morita-equivalent.*

Proof The first statement is immediate from combining the last four lemmas. The second is simply its translation into the language of coherent theories. □

We may also state the theorem in a purely topos-theoretic form, as follows:

Corollary 3.5.10 *Let $\mathcal{E}$ and $\mathcal{F}$ be coherent toposes, and let $f: \mathcal{F} \to \mathcal{E}$ be a geometric morphism such that f^* maps coherent objects of $\mathcal{E}$ to coherent objects of $\mathcal{F}$, and such that composition with f induces a weak equivalence $\mathfrak{Top}/\mathbf{Set}(\mathbf{Set}, \mathcal{F}) \to \mathfrak{Top}/\mathbf{Set}(\mathbf{Set}, \mathcal{E})$. Then f is an equivalence.*

Proof Since f^* preserves coherent objects, it restricts to a coherent functor between the pretoposes of coherent objects of $\mathcal{E}$ and $\mathcal{F}$; the second hypothesis on f implies that this functor satisfies the hypothesis of 3.5.9(i). □

As an application of conceptual completeness (in fact of 3.5.7), we show that the situation encountered in Example 3.1.15 cannot occur amongst coherent toposes.

Corollary 3.5.11

(i) *Let $\mathcal{E}$ be a coherent topos such that the category $\mathfrak{Top}/\mathbf{Set}(\mathbf{Set}, \mathcal{E})$ is a preorder. Then $\mathcal{E}$ is localic over $\mathbf{Set}$.*

(ii) *Let* $\mathbb{T}$ *be a coherent theory such that the category* $\mathbb{T}\text{-Mod}(\mathbf{Set})$ *is a preorder. Then* $\mathbb{T}$ *is Morita-equivalent to a propositional theory.*

Proof The two statements are clearly equivalent; we shall prove the first. Suppose $\mathcal{E} \simeq \mathbf{Sh}(\mathcal{C}, T)$ where $\mathcal{C}$ is a small pretopos and T is its coherent coverage; then, as we saw in 3.3.15, we may identify its localic reflection $\mathcal{L}$ (over **Set**) with $\mathbf{Sh}(\mathrm{Sub}_{\mathcal{C}}(1), T')$ where T' is the coherent coverage on $\mathrm{Sub}_{\mathcal{C}}(1)$. Moreover, the hyperconnected geometric morphism $h \colon \mathcal{E} \to \mathcal{L}$ may be identified with that induced by the inclusion functor $I \colon \mathrm{Sub}_{\mathcal{C}}(1) \to \mathcal{C}$, as we saw in C2.3.22. Now I is clearly conservative and full on subobjects; and the functor $\mathfrak{Coh}(\mathcal{C}, \mathbf{Set}) \to \mathfrak{Coh}(\mathrm{Sub}_{\mathcal{C}}(1), \mathbf{Set})$ which it induces is trivially faithful, since its domain is a preorder. So by 3.5.7 I is finitely covering. It is now easy to see that I induces a (weak) equivalence from the pretopos completion of $\mathrm{Sub}_{\mathcal{C}}(1)$ to $\mathcal{C}$; hence h is an equivalence, and $\mathcal{E}$ is localic. □

Remark 3.5.12 As we mentioned earlier, the definability theorem 3.5.1 does not hold (in the form in which we stated it) for regular logic. Nevertheless, the conceptual completeness theorem does hold: *if* $\mathcal{C}$ *and* $\mathcal{D}$ *are effective regular categories, and* $I \colon \mathcal{C} \to \mathcal{D}$ *is a regular functor such that* $I^* \colon \mathfrak{Reg}(\mathcal{D}, \mathbf{Set}) \to \mathfrak{Reg}(\mathcal{C}, \mathbf{Set})$ *is a weak equivalence, then* I *is a weak equivalence.* The key ingredient which we need, in order to adapt the proof just given for coherent logic to the regular case, is Lemma 3.3.11. For example, in the proof of 3.5.6, the definability theorem yields, for a subobject $m \colon B \rightarrowtail IA$ in $\mathcal{D}$, a coherent formula-in-context $x \,.\, \phi$ (where x is of sort A) over the signature of $\mathbb{T}_{\mathcal{C}}$, such that the regular formula $x \,.\, (\exists y)(my = x)$ is provably equivalent in $\mathbb{T}_{\mathcal{D}}$ to (the translation via I^* of) $x \,.\, \phi$. But we may write ϕ as a finite disjunction $\bigvee_{i=1}^{n} \phi_i$ of regular formulae, by 1.3.8(ii); so in the case when $\mathbb{T}_{\mathcal{D}}$ is a regular theory we may conclude from 3.3.11 that $x \,.\, (\exists y)(my = x)$ is actually equivalent to one of the $x \,.\, \phi_i$. Similar modifications need to be made to the proof of 3.5.7 (in which we replace the finite covering family of partial maps by a single partial map); we omit the details.

Remark 3.5.13 We should also mention that a substantially different proof of the conceptual completeness theorem has been given by A. M. Pitts [972]. Pitts's proof is more categorical in spirit, and derives the result from an 'interpolation property' for cocomma squares in the 2-category of pretoposes. In [974], Pitts extended this result to the 2-category of Heyting pretoposes (that is, pretoposes which are Heyting categories) and obtained a conceptual completeness theorem for full first-order logic. Unfortunately, the amount of machinery which has to be developed on the way to Pitts's proof is such as to place it beyond the scope of this book.

Suggestions for further reading: Ballard & Boshuck [43], Butz & Moerdijk [216], Makkai [781], Makkai & Reyes [790], Pitts [972–974], Zawadowski [1260].

D4

HIGHER-ORDER LOGIC

D4.1 Interpreting higher-order logic in a topos

We now turn from first-order logic to higher-order logic: the principal extra feature of this logic, compared with that which we studied in Chapter D1, is that we have variables running over types whose intended interpretation is the set of all subsets of (the set denoted by) some other type, or the set of all functions between two such types – and we can therefore quantify over elements of such 'higher types'. The extra feature of our formal languages which makes this possible is the presence of *type constructors*, which build new types from old. In fact, as we remarked after Definition 1.1.1, the first of these type constructors (the product type constructor) was implicitly present already in our account of first-order logic, in that our function symbols and relation symbols had types assigned to them which were finite strings of sorts, and which were formally interpreted as products of those sorts. Now that we need other type constructors, it makes sense to add the product type constructor explicitly to our definition of a signature. We shall find it convenient to maintain the distinction between *sorts* which are the basic types given to us, along with the primitive function symbols and relation symbols, as part of the specification of our signature, and the *types* of the signature which may be either sorts or 'compound types' built up using type constructors.

Definition 4.1.1 A *signature* Σ for higher-order logic is defined by specifying a set Σ-Sort of *sorts*, a set Σ-Fun of *function symbols* and a set Σ-Rel of *relation symbols* as in 1.1.1, together with a number of *type constructors* chosen from the following list, which recursively define the set Σ-Typ of *types*:

(i) *Basic types*: Each sort is a type.

(ii) *Product types*: There is a distinguished type 1 (the empty product); and if A and B are types, then $A \times B$ is a type.

(iii) *Function types*: If A and B are types, then $[A \to B]$ is a type.

(iv) *Power types*: If A is a type, then PA is a type.

(v) *List types*: If A is a type, then LA is a type.

To each function symbol $f \in \Sigma$-Fun we assign an ordered pair (A, B) of elements of Σ-Typ; as before, we write $f : A \to B$ to indicate that f has the pair of types

(A, B). Similarly, each relation symbol $R \in \Sigma$-Rel has a type in Σ-Typ, and we write $R \rightarrowtail A$ to indicate that R has type A.

An alternative approach (which is used in [682], for example) is to assume given the entire set Σ-Typ as part of the primitive data, together with binary operations $(A, B \mapsto A \times B)$, $(A, B \mapsto [A \rightarrow B])$ and so on, on this set. Clearly, Definition 4.1.1 is more restrictive, in that we require Σ-Typ to be *freely* generated by the type constructors from the subset Σ-Sort of basic types; however, there are occasions when it is useful to be able to argue by induction over the 'complexity' of types, and for this purpose the recursive definition of types which we have given is essential.

If our signature has the type constructors (ii) and (iv), we shall often write Ω for the particular type $P1$. Similarly, if it has constructors (ii) and (v), we write N (the type of natural numbers) for $L1$. Sometimes we may wish to consider signatures which have the type N but not the general list type constructor. From now on, all our signatures will contain the product type constructor; thus we have restricted all our primitive function and relation symbols to be unary. (If our signature also contains function types, then we can further restrict our function symbols to be constants (function symbols of type $(1, B)$ for some B), since a function symbol $f \colon A \rightarrow B$ can be replaced by $\overline{f} \colon 1 \rightarrow [A \rightarrow B]$. And if Σ has power types, then we can replace all primitive relation symbols by constants as well: a relation $R \rightarrowtail A$ becomes a constant $\ulcorner R \urcorner \colon 1 \rightarrow PA$.)

Associated with each type constructor, we have term constructors enabling us to build terms of the various compound types. The term constructor for power types in fact builds terms from formulae, so over a signature which contains power types we have to define terms and formulae by a simultaneous recursion; nevertheless, we shall present the recursive clauses of the definition for terms and for formulae separately in the two definitions which follow. Another feature which was not present in first-order logic is that some of the term constructors have the effect of binding variables, and so we have to keep track of the set $\mathrm{FV}(t)$ of free variables in a term t, as well as the free variables in a formula.

Definition 4.1.2 Let Σ be a higher-order signature. The collection of *terms* over Σ is defined recursively, together with the *type* of each term (we write $t \colon A$ to indicate that the term t has type A) and the (finite) set $\mathrm{FV}(t)$ of *free variables* of each term t, by the following clauses:

(i) We have a stock of *variables* $x \colon A$ for each type A; $\mathrm{FV}(x) = \{x\}$.

(ii) If $f \colon A \rightarrow B$ is a function symbol of Σ and $t \colon A$, then $f(t) \colon B$; $\mathrm{FV}(f(t)) = \mathrm{FV}(t)$. (Since we have now restricted ourselves to unary function symbols, we often write simply ft rather than $f(t)$.)

(iii) (*a*) We have a distinguished term $* \colon 1$; $\mathrm{FV}(*) = \emptyset$.

(*b*) If $s \colon A$ and $t \colon B$, then $\langle s, t \rangle \colon A \times B$; $\mathrm{FV}(\langle s, t \rangle) = \mathrm{FV}(s) \cup \mathrm{FV}(t)$.

(iv) If $t\colon A \times B$, then $\mathsf{fst}(t) : A$ and $\mathsf{snd}(t) : B$; both of these terms have the same free variables as t.

(v) If $t : B$ and x is a variable of type A, then $\lambda x\colon A.t : [A \to B]$; $\mathrm{FV}(\lambda x\colon A.t) = \mathrm{FV}(t) \setminus \{x\}$. (If we do not need to specify the type of the variable x, then we simply write $\lambda x.t$ for $\lambda x\colon A.t$.)

(vi) If $s\colon [A \to B]$ and $t\colon A$, then $\mathsf{app}(s,t)\colon B$; $\mathrm{FV}(\mathsf{app}(s,t)) = \mathrm{FV}(s) \cup \mathrm{FV}(t)$. (We often abbreviate $\mathsf{app}(s,t)$ to $s(t)$, or even to st, but for formal purposes it is useful to have a notation which distinguishes between application of a function symbol to a term, as in (ii), and application of a term whose type is a function type.)

(vii) If ϕ is a formula and x is a variable of type A, then $\{x : A \mid \phi\} : PA$; $\mathrm{FV}(\{x\colon A \mid \phi\}) = \mathrm{FV}(\phi) \setminus \{x\}$. (Again, we may abbreviate this to $\{x \mid \phi\}$ if it is not necessary to specify the type of x.)

(viii) (a) For each type A, there is a distinguished term $[\,]$ of type LA; $\mathrm{FV}([\,]) = \emptyset$.

(b) If $s : A$ and $t : LA$, then $\mathsf{cons}(s,t) : LA$; $\mathrm{FV}(\mathsf{cons}(s,t)) = \mathrm{FV}(s) \cup \mathrm{FV}(t)$. (Intuitively, $\mathsf{cons}(s,t)$ is the list obtained by appending s to the list t. In the particular case $A = 1$, we normally write $\mathsf{succ}(t)$ for $\mathsf{cons}(*,t)$ and zero for $[\,]$.)

(ix) The remaining term constructor associated with list types, the *iterator*, is more complicated. Suppose we have variables x, y of types A, B respectively (here A and B may be the same type, but x and y must be distinct), and suppose given terms s, t, u of types A, A, LB respectively, such that x and y are not free in either t or u. Then $\mathsf{iter}_{x,y}(s,t,u)$ is a term of type A, in which all occurrences of x and y in s are bound, but the other free variables of s, t and u remain free. (The intuition behind this definition is that A denotes a set of 'values' and B a set of 'instructions': the term s (which is implicitly assumed to involve the variables x and y, even though it need not do so in practice) is thought of as a 'machine' which, given an input value x and an instruction y, produces an output value from them. Then $\mathsf{iter}_{x,y}(s,t,u)$ denotes the value at which we arrive, starting from the input value t, if we make the machine follow the 'program' (list of instructions) u.) In the particular case $B = 1$ (so that $LB = N$), we can simplify this constructor by omitting the variable y: thus we simply write $\mathsf{iter}_x(s,t,u)$ in this case.

Of course, clauses (v) and (vi) of the definition apply only if we have function types in Σ, clause (vii) only if we have power types, and clauses (viii) and (ix) only if we have list types. Note that, except for power types, the clauses come in pairs consisting of one or more 'constructors' which build terms of a compound type from terms of simpler types (clauses (iii), (v), (viii)) and one or more 'destructors' which build terms of simpler types from terms of the compound type (clauses (iv), (vi) and (ix)). The reason why we have no 'destructor' for

power types is that, just as the constructor for power types builds terms out of formulae rather than terms, the corresponding destructor builds formulae out of terms of a power type – it thus appears as clause (iii) of the next definition.

Definition 4.1.3 The *atomic formulae* over Σ are of three kinds:

(i) If $R \rightarrowtail A$ is a relation symbol and $t\colon A$, then $R(t)$ is an atomic formula; $\mathrm{FV}(R(t)) = \mathrm{FV}(t)$.

(ii) If s and t are terms of the same type A, then $(s =_A t)$ is an atomic formula; $\mathrm{FV}(s =_A t) = \mathrm{FV}(s) \cup \mathrm{FV}(t)$.

(iii) If $s : A$ and $t : PA$ for some A, then $(s \in_A t)$ is an atomic formula; $\mathrm{FV}(s \in_A t) = \mathrm{FV}(s) \cup \mathrm{FV}(t)$.

The compound formulae over Σ are built up from the atomic formulae exactly as in clauses (iii)–(x) of Definition 1.1.3.

We normally omit the type subscripts from $=_A$ and $\in_A$ whenever it is safe to do so.

In order to define the notion of Σ-structure in a category $\mathcal{C}$, we need to know that $\mathcal{C}$ has the categorical structure required to interpret the type constructors present in Σ. Thus, if Σ contains product types, we require $\mathcal{C}$ to have finite products; if Σ contains function types, we require $\mathcal{C}$ to be cartesian closed; if Σ contains power types, we require $\mathcal{C}$ to have power objects (i.e. to be a topos), and if Σ contains list types, then we require $\mathcal{C}$ to have list objects as defined in A2.5.15. (In fact if $\mathcal{C}$ is not cartesian closed, then we need to strengthen the definition of list object along the lines suggested in A2.5.3: we say that (LA, o_A, s_A) is a list object over A if, given morphisms $b\colon C \to B$ and $t\colon A \times B \times C \to B$, there is a unique $f\colon LA \times C \to B$ making

$$\begin{tikzcd}
1 \times C \arrow[r, "o_A \times 1"] \arrow[d, "\pi_2"] & LA \times C \arrow[d, "f"] & A \times LA \times C \arrow[l, "s_A \times 1"'] \arrow[d, "(1 \times f, \pi_3)"] \\
C \arrow[r, "b"] & B & A \times B \times C \arrow[l, "t"']
\end{tikzcd}$$

commute. If $\mathcal{C}$ is cartesian closed, then this version of the definition may easily be deduced from the special case $C = 1$ which we gave as the definition in A2.5.15, just as we deduced A2.5.2 from A2.5.1 for natural number objects.) We shall summarize the above restrictions on $\mathcal{C}$ by saying '$\mathcal{C}$ is a suitable category for Σ'.

Definition 4.1.4 Let Σ be a higher-order signature, and $\mathcal{C}$ a suitable category for Σ. A Σ-*structure* M in $\mathcal{C}$ is defined by specifying a function Σ-Sort $\to \mathrm{ob}\,\mathcal{C}$, which we denote $A \mapsto MA$ and which we extend to a function Σ-Typ $\to \mathrm{ob}\,\mathcal{C}$ 'in the obvious way'; i.e. we define $M1$ to be the terminal object 1 of $\mathcal{C}$, $M(A \times B)$ to be $MA \times MB$, $M[A \to B]$ to be the exponential MB^{MA}, $M(PA)$ to be the

power object $P(MA)$ and $M(LA)$ to be the list object $L(MA)$. In addition, M is required to specify a morphism $Mf\colon MA \to MB$ for each function symbol $f\colon A \to B$ of Σ, and a subobject $MR \rightarrowtail MA$ for each relation symbol $R \rightarrowtail A$ of Σ.

Note that, in contrast to Definition 1.2.1, we have *not* defined the notion of a homomorphism of Σ-structures. The reason is that the power object functor is contravariant, and the exponential functor $(A, B) \mapsto B^A$ is contravariant in its first variable; so if we are given a family of morphisms $MA \to NA$ for each sort A, we will not in general obtain morphisms $MA \to NA$ for each type A. Nevertheless, we can define a notion of *isomorphism* of Σ-structures in an obvious way; and these are the only morphisms of higher-order structures which we shall need to consider in practice.

The notions of *context*, of *term-in-context*, of *formula-in-context* and of *sequent* are all defined just as they were in the first-order case.

Definition 4.1.5 If M is a Σ-structure in a suitable category $\mathcal{C}$ (and in addition $\mathcal{C}$ has the structure required to interpret the propositional connectives and quantifiers involved), we assign to each term-in-context $\vec{x}\,.\,t$ of type B an interpretation $[\![\vec{x}\,.\,t]\!]_M\colon MA_1 \times \cdots \times MA_n \to MB$ (where $A_1, \ldots, A_n$ is the string of types of the variables in the context $\vec{x}$), and to each formula-in-context $\vec{x}\,.\,\phi$ an interpretation $[\![\vec{x}\,.\,\phi]\!]_M \rightarrowtail MA_1 \times \cdots \times MA_n$, by the following clauses (in which we omit the subscript M's, and write A for the product type $A_1 \times \cdots \times A_n$):

(i) If t is a variable x_i, then $[\![\vec{x}\,.\,t]\!]$ is the projection π_i.

(ii) If t is $f(s)$ (where $s\colon C$, say), then $[\![\vec{x}\,.\,t]\!]$ is the composite

$$MA \xrightarrow{[\![\vec{x}\,.\,s]\!]} MC \xrightarrow{Mf} MB\,.$$

(iii) If t is $*$, then $[\![\vec{x}\,.\,t]\!]$ is the unique morphism $MA \to 1$.

(iv) If $B = B_1 \times B_2$ and t is $\langle r, s\rangle$, then $[\![\vec{x}\,.\,t]\!]$ is

$$MA \xrightarrow{([\![\vec{x}\,.\,r]\!],\,[\![\vec{x}\,.\,s]\!])} MB_1 \times MB_2\,.$$

(v) If t is $\mathsf{fst}(s)$ where $u\colon B \times C$, then $[\![\vec{x}\,.\,t]\!]$ is

$$MA \xrightarrow{[\![\vec{x}\,.\,s]\!]} MB \times MC \xrightarrow{\pi_1} MB$$

(and similarly if t is $\mathsf{snd}(s)$).

(vi) If $B = [C \to D]$ and t is $\lambda z\,.\,s$ (where we assume for convenience – by renaming the bound variable if necessary – that z does not appear in the context $\vec{x}$), then $[\![\vec{x}\,.\,t]\!]$ is the exponential transpose of

$$MA \times MC \xrightarrow{[\![\vec{x}, z\,.\,s]\!]} MD\,.$$

(vii) If t is $\mathsf{app}(r, s)$ where $r\colon [C \to B]$, then $[\![\vec{x}\,.\,t]\!]$ is

$$MA \xrightarrow{([\![\vec{x}\,.\,r]\!],[\![\vec{x}\,.\,s]\!])} MB^{MC} \times MC \xrightarrow{\mathrm{ev}} MB\,.$$

(viii) If $B = PC$ and t is $\{z\colon C \mid \phi\}$ (where, once again, we assume that z does not appear in $\vec{x}$), then $[\![\vec{x}\,.\,t]\!]$ is the name of the relation

$$[\![\vec{x}, z\,.\,\phi]\!] \rightarrowtail MA \times MC\,.$$

(ix) If $B = LC$ and t is $[\,]$, then $[\![\vec{x}\,.\,t]\!]$ is

$$MA \longrightarrow 1 \xrightarrow{o_{MC}} L(MC)\,.$$

(x) Similarly, if $B = LC$ and t is $\mathsf{cons}(r, s)$, then $[\![\vec{x}\,.\,t]\!]$ is

$$MA \xrightarrow{([\![\vec{x}\,.\,r]\!],[\![\vec{x}\,.\,s]\!])} MC \times L(MC) \xrightarrow{s_{MC}} L(MC)\,.$$

(xi) If t is $\mathsf{iter}_{y,z}(r, s, u)$ where $y\colon B$ and $z\colon C$, then $[\![\vec{x}\,.\,t]\!]$ is the composite

$$MA \xrightarrow{([\![\vec{x}\,.\,u]\!],1)} L(MC) \times MA \xrightarrow{f} MB,$$

where f is the unique morphism making

$$\begin{tikzcd}
1 \times MA \arrow[r, "o_{MC} \times 1"] \arrow[d, "\pi_2"] & L(MC) \times MA \arrow[d, "f"] & MC \times L(MC) \times MA \arrow[l, "s_{MC} \times 1"'] \arrow[d, "(1 \times f, \pi_3)"] \\
MA \arrow[r, "{[\![\vec{x}\,.\,s]\!]}"] & MB & MC \times MB \times MA \arrow[l, "{[\![z, y, \vec{x}\,.\,r]\!]}"']
\end{tikzcd}$$

commute.

(xii) If ϕ is the atomic formula $(s \in_B t)$, then $[\![\vec{x}\,.\,\phi]\!] \rightarrowtail MA$ is the subobject obtained by pulling back the universal relation $\in_{MB} \rightarrowtail MB \times P(MB)$ along

$$MA \xrightarrow{([\![\vec{x}\,.\,s]\!],[\![\vec{x}\,.\,t]\!])} MB \times P(MB)\,.$$

(xii–xxiii) The interpretations of the other atomic formulae, and of all compound formulae, are defined exactly as in clauses (i)–(x) of Definition 1.2.6.

Of course, we say a Σ-structure M *satisfies* a sequent $(\phi \vdash_{\vec{x}} \psi)$ if $[\![\vec{x}\,.\,\phi]\!] \leq [\![\vec{x}\,.\,\psi]\!]$ in $\mathrm{Sub}(MA)$; and we say M is a *model* for a theory $\mathbb{T}$ over Σ (that is, a set of sequents over Σ) if it satisfies all the sequents in $\mathbb{T}$.

As in the first-order case, we have a Substitution Property for the terms and formulae of our higher-order languages, which is proved in the same way as 1.2.4 and 1.2.7.

Lemma 4.1.6 *Suppose M is a Σ-structure in a suitable category $\mathcal{C}$ (and that the first-order structure of $\mathcal{C}$, to the extent that it is used in the interpretation of the terms and formulae mentioned below, is stable under pullback). Let $\vec{y}\,.\,t$ be a term-in-context (respectively, let $\vec{y}\,.\,\phi$ be a formula-in-context) over the signature Σ; let $\vec{s}$ be a string of terms of the same length and type as the string $\vec{y}$, and let $\vec{x}$ be a context suitable for all the terms in $\vec{s}$. Then $[\![\vec{x}\,.\,t[\vec{s}/\vec{y}]]\!]$ is the composite*

$$M(A_1 \times \cdots \times A_n) \xrightarrow{([\![\vec{x}\,.\,s_1]\!], \ldots, [\![\vec{x}\,.\,s_m]\!])} M(B_1 \times \cdots \times B_m) \xrightarrow{[\![\vec{y}\,.\,t]\!]} MC$$

(respectively, $[\![\vec{x}\,.\,\phi[\vec{s}/\vec{y}]]\!]$ is the pullback of $[\![\vec{y}\,.\,\phi]\!] \rightarrowtail M(B_1 \times \cdots \times B_m)$ along the first factor of the above composite), where A_i, B_j and C are the types of x_i, y_j and t respectively. □

We next record the relationships between the term constructors and destructors introduced in 4.1.2 and 4.1.3, which express in the formal language the fact that they are associated with certain bijections in the categorical structure. We say a sequent is *valid* in a (suitable) category $\mathcal{C}$ if it is satisfied by every structure (for the appropriate signature) in $\mathcal{C}$.

Lemma 4.1.7

(i) *In any category with finite products, the sequents $(\top \vdash_w (w = *))$,*

$$(\top \vdash_{x,y} (\mathsf{fst}(\langle x, y\rangle) = x)), \quad (\top \vdash_{x,y} (\mathsf{snd}(\langle x, y\rangle) = y)) \quad \textit{and}$$
$$(\top \vdash_z (\langle \mathsf{fst}(z), \mathsf{snd}(z)\rangle = z))$$

are valid. (Here w and z are variables of types 1 and $A \times B$ respectively; the types of x and y are arbitrary.)

(ii) *In any cartesian closed category, the sequents*

$$(\top \vdash_{\vec{x},z} (\mathsf{app}(\lambda y\,.\,t, z) = t[z/y])) \quad \textit{and} \quad (\top \vdash_u (\lambda y\,.\,\mathsf{app}(u, y) = u))$$

are valid. (Here t is a term of arbitrary type whose free variables are in the string $\vec{x}, y$, and u is a variable of function type.)

(iii) *In any topos, the sequents*

$$((z \in_A \{y\colon A \mid \phi\}) \dashv\vdash_{\vec{x},z} \phi[z/y]) \quad \textit{and} \quad (\top \vdash_w (\{x\colon A \mid x \in_A w\} = w))$$

are valid. (Here ϕ is any formula with free variables in the string $\vec{x}, y$, and w is a variable of type PA.)

(iv) *In any category with finite products and list objects, the three sequents* $(\top \vdash_{\vec{z},w} (\mathsf{iter}_{x,y}(s, w, [\,]) = w))$,

$$(\top \vdash_{\vec{z},u,v,w} (\mathsf{iter}_{x,y}(s, w, \mathsf{cons}(u, v)) = s[\mathsf{iter}_{x,y}(s, w, v), u/x, y]))$$

and $(\top \vdash_z (\mathsf{iter}_{x,y}(\mathsf{cons}(y, x), [\,], z) = z))$ *are valid. (In the first two of these, s is a term whose free variables are in the string $x, y, \vec{z}$, and in the third x, y, z are variables of types LA, A, LA respectively.)*

Proof These are all straightforward when we translate them into assertions about morphisms in categories via the clauses of 4.1.5. For example, the two sequents in (ii) assert that transpositions in the two directions across the exponential adjunction are inverse to each other, and (iii) similarly encodes the bijection between morphisms $A \to PB$ in a topos and subobjects of $A \times B$. The first two sequents of (iv) are just the commutativity of the two cells in the diagram of 4.1.5(xi), and the third one follows from the fact that the unique morphism $LA \to LA$ commuting with o_A and s_A is the identity. □

As in Section D1.3, we may associate with any suitable category $\mathcal{C}$ a *canonical signature* $\Sigma_{\mathcal{C}}$, with one sort $\ulcorner A \urcorner$ for each object A of $\mathcal{C}$, one function symbol $\ulcorner f \urcorner : \ulcorner A \urcorner \to \ulcorner B \urcorner$ for each morphism $f : A \to B$ of $\mathcal{C}$, and one relation symbol $\ulcorner R \urcorner \rightarrowtail \ulcorner A \urcorner$ for each monomorphism $R \rightarrowtail A$ in $\mathcal{C}$. (Strictly speaking, we should require $\mathcal{C}$ to be small here, but in practice we shall use this notation informally for non-small categories as well. Provided $\mathcal{C}$ has the categorical structure needed to interpret the type constructors of $\Sigma_{\mathcal{C}}$, we then have a canonical $\Sigma_{\mathcal{C}}$-structure in $\mathcal{C}$ (to which we do not bother to give a name).)

Note, incidentally, that $\Sigma_{\mathcal{C}}$ has many more types than $\mathcal{C}$ has objects, since $\ulcorner A \urcorner \times \ulcorner B \urcorner$ is different from $\ulcorner A \times B \urcorner$, and so on, although they are interpreted as the same object of $\mathcal{C}$. Informally, we shall tend to identify two types of $\Sigma_{\mathcal{C}}$ if they 'become equal on removing the $\ulcorner\ \urcorner$'s', and thus if (for example) $f : PA \to B^C$ is a morphism of $\mathcal{C}$ we shall feel free to use $\ulcorner f \urcorner$ as a function symbol of type $P\ulcorner A \urcorner \to [\ulcorner C \urcorner \to \ulcorner B \urcorner]$, rather than $\ulcorner PA \urcorner \to \ulcorner B^C \urcorner$, if it seems appropriate to do so. Similarly, we shall feel free to commit such abuses of notation as writing $\{\langle x, y \rangle : A \times B \mid \phi\}$, where ϕ is a formula having free variables x and y of types A and B, for the term of type $P(\ulcorner A \urcorner \times \ulcorner B \urcorner)$ (or of type $\ulcorner P(A \times B) \urcorner$) which should properly be written as $\{w : \ulcorner A \urcorner \times \ulcorner B \urcorner \mid \phi[\mathsf{fst}(w), \mathsf{snd}(w)/x, y]\}$. For the formal mechanism needed to make this precise, see 4.2.7 below.

Of course, we say that $\mathcal{C}$ *satisfies* a higher-order sequent σ over $\Sigma_{\mathcal{C}}$, and write $\mathcal{C} \models \sigma$, if the nameless $\Sigma_{\mathcal{C}}$-structure in $\mathcal{C}$ satisfies σ. (In the particular case – which in fact suffices when we have full first-order logic available, as we remarked after 1.1.5 – that σ is of the form $(\top \vdash_{[\,]} \phi)$, we shall simply say that $\mathcal{C}$ satisfies the sentence ϕ.) Thus we may use the higher-order language over $\Sigma_{\mathcal{C}}$ to reason about objects and morphisms of $\mathcal{C}$, as we did in the first-order case in 1.3.13 and 1.3.14. In the particular case when $\mathcal{C}$ is a topos, we shall often refer to the higher-order language over $\Sigma_{\mathcal{C}}$ as the *internal language of* $\mathcal{C}$.

As an example of the uses to which the canonical signature may be put, we prove a long-promised result, the 'Cantor diagonal argument' for toposes.

Proposition 4.1.8

(i) *If a topos $\mathcal{E}$ contains an object A such that PA is a subquotient of A (cf.* A4.6.1*), then $\mathcal{E}$ is degenerate (i.e. equivalent to the terminal category* **1***).*

(ii) *If there exists a monomorphism $A^A \rightarrowtail A$ in a Boolean topos $\mathcal{E}$, then $A \cong 1$.*

Proof (i) Note first that if we are given morphisms

$$PA \xleftarrow{\;g\;} B \xrightarrow{\;f\;} A$$

expressing PA as a subquotient of A, then on applying the covariant and contravariant power-object functors we obtain

$$\begin{tikzcd} PPA \arrow[r, shift left, "Pg"] & PB \arrow[l, shift left, "\exists g"] \arrow[r, shift left, "\exists f"] & PA \arrow[l, shift left, "Pf"] \end{tikzcd}$$

expressing PPA as a retract of PA, by A2.2.5 and A2.3.6(iii); so we might as well start from the assumption that we have morphisms $i\colon PA \rightarrowtail A$ and $r\colon A \twoheadrightarrow PA$ with $ri = 1_{PA}$. Now consider the closed term

$$\{x\colon A \mid \neg(x \in_A \ulcorner r \urcorner(x))\}$$

of type PA; let us denote this term by c (for Cantor). Is the sentence $\phi \equiv (\ulcorner i \urcorner(c) \in_A c)$ satisfied in $\mathcal{E}$? If so, then we have $\mathcal{E} \models \neg(\ulcorner i \urcorner(c) \in_A \ulcorner r \urcorner(\ulcorner i \urcorner(c)))$; but we also have $\mathcal{E} \models (\ulcorner r \urcorner(\ulcorner i \urcorner(c)) = c)$, and so we deduce that $\mathcal{E} \models (\phi \vdash \neg\phi)$, whence $\mathcal{E} \models (\phi \vdash \bot)$, i.e. $\mathcal{E} \models \neg\phi$. But from $\neg\phi$ we can similarly deduce $\neg\neg\phi$, and hence $\bot$, so $\mathcal{E} \models \bot$; but the latter is exactly the statement that $\mathcal{E}$ is degenerate.

(ii) The hypothesis clearly implies that we have a morphism $1 \to A$, since there exists a morphism $\overline{1_A}\colon 1 \to A^A$; so we need only prove that $A \to 1$ is monic, i.e. that the sentence $(\forall x, y\colon A)(x = y)$ is satisfied in $\mathcal{E}$. But since $\mathcal{E}$ is Boolean, it satisfies

$$((\forall x, y\colon A)(x = y) \vee (\exists x, y\colon A)\neg(x = y)) \;;$$

let $U \rightarrowtail 1$ and $V \rightarrowtail 1$ be the interpretations of the two halves of this disjunction. Interpreting the second half, we see that there is an epimorphism $B \twoheadrightarrow V$ and a monomorphism $B^*(2) \rightarrowtail B^*(A)$ in $\mathcal{E}/B$. But $B^*(2)$ is the subobject classifier of the latter topos, and so we obtain a monomorphism $P(B^*(A)) \cong B^*(2)^{B^*(A)} \rightarrowtail B^*(A)^{B^*(A)} \cong B^*(A^A) \rightarrowtail B^*(A)$. Thus by part (i) $\mathcal{E}/B$ is degenerate, so $B \cong 0$; hence $V \cong 0$ and $U \cong 1$, as required. □

The assumption of Booleanness in part (ii) of 4.1.8 cannot be omitted. We shall see in 4.2.16 below that there exist cartesian closed categories $\mathcal{C}$ (which may be taken to be small) containing nontrivial examples of objects A such that $A \cong A^A$; and by A1.5.6(i) these examples are preserved by the Yoneda embedding $\mathcal{C} \to [\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$. The following example is also of interest:

Example 4.1.9 *There exists a non-degenerate (non-Boolean) topos $\mathcal{E}$ with a natural number object N, such that N^N is a subquotient of N.* We construct $\mathcal{E}$ as the classifying topos $\mathbf{Set}[\mathbb{T}]$ of a propositional geometric theory $\mathbb{T}$, namely the theory of 'partial surjections from $\mathbb{N}$ to $\mathbb{N}^{\mathbb{N}}$', which is (the case $A = \mathbb{N}^{\mathbb{N}}$ of) the intersection of the two theories considered in C1.2.8 and C1.2.9. Explicitly, we take a signature with primitive propositions $p(n, f)$ indexed by pairs consisting of a natural number n and a function $f \colon \mathbb{N} \to \mathbb{N}$, with the axioms

$$((p(n,f) \wedge p(n,g)) \vdash \bot)$$

whenever $f \neq g$, and

$$(\top \vdash \bigvee \{p(n,f) \mid n \in \mathbb{N}\})$$

for each $f \in \mathbb{N}^{\mathbb{N}}$. We shall also write $\mathbb{T}_0$ for the subtheory of $\mathbb{T}$ having only the first group of axioms, i.e. the theory of partial maps $\mathbb{N} \rightharpoonup \mathbb{N}^{\mathbb{N}}$.

It is easy to see that the classifying topos of $\mathbb{T}_0$ may be identified with the functor category $[P^{\mathrm{op}}, \mathbf{Set}]$, where P is the poset of finite conjunctions $\bigwedge_{i=1}^{n} p(n_i, f_i)$ which are consistent (i.e. satisfy $n_i \neq n_j$ whenever $i \neq j$), ordered by entailment. Equivalently, we may think of P as the poset of partial functions $\mathbb{N} \rightharpoonup \mathbb{N}^{\mathbb{N}}$ with finite domain, ordered by the relation of extension, as in A2.1.11(g). Moreover, the coverage on this poset generated by the second group of axioms is exactly that considered in A2.1.11(g): explicitly, for each $f \colon \mathbb{N} \to \mathbb{N}$, each finite partial function $\phi \colon \mathbb{N} \rightharpoonup \mathbb{N}^{\mathbb{N}}$ is covered by the set of all ϕ' which extend ϕ and have f in their image. We note that these covers are all inhabited, so $\mathcal{E} = \mathbf{Set}[\mathbb{T}] \simeq \mathbf{Sh}(P, T)$ is a dense subtopos of $[P^{\mathrm{op}}, \mathbf{Set}] \simeq \mathbf{Set}[\mathbb{T}_0]$, and in particular non-degenerate. Moreover, the generic $\mathbb{T}$-model expresses $\gamma^*(\mathbb{N}^{\mathbb{N}})$ as a subquotient of $\gamma^*(\mathbb{N})$ in $\mathcal{E}$, where $\gamma \colon \mathcal{E} \to \mathbf{Set}$ is the unique geometric morphism; and we know that $\gamma^*(\mathbb{N})$ is the natural number object of $\mathcal{E}$, by A4.1.14. But we also know that the exponential N^N in $[P^{\mathrm{op}}, \mathbf{Set}]$ is simply the constant functor with value $\mathbb{N}^{\mathbb{N}}$, by A1.5.6(ii); and this constant functor is a sheaf by A2.1.11(g), and is therefore $\gamma^*(\mathbb{N}^{\mathbb{N}})$. Moreover, the inclusion $\mathbf{Sh}(P,T) \to [P^{\mathrm{op}}, \mathbf{Set}]$ is a cartesian closed functor by A4.2.9, so this object must also be the exponential N^N as computed in the subtopos. (Alternatively, we could argue using C3.3.11(c) that $\mathcal{E}$ is a locally connected topos, and so γ^* preserves exponentials.)

We remark that the effective topos (which we shall study in Section F2.2) provides another example of a topos in which N^N is a subquotient of N; thus the main interest of 4.1.9 is in the fact that this phenomenon can occur in a Grothendieck topos. Note also that if we tried to replace the theory of partial

surjections by the theory of injections $\mathbb{N}^{\mathbb{N}} \rightarrowtail \mathbb{N}$, as in C1.2.9 (that is, if we added the axioms $(\top \vdash \bigvee\{p(n,f) \mid f \in \mathbb{N}^{\mathbb{N}}\})$ for each $n \in \mathbb{N}$), we would not obtain a locally connected topos; and in fact the exponential N^N in this topos would be bigger than $\gamma^*(\mathbb{N}^{\mathbb{N}})$ (and so would not occur as a subobject of N). Intuitively, this is because the presence of a generic monomorphism $f: \mathbb{N}^{\mathbb{N}} \rightarrowtail \mathbb{N}$ allows us to define 'new' functions $N \to N$ which were not present in our original topos of sets (for example, the function whose value at n is f applied to the constant function with value n), whereas the partial surjection which is present in the topos of 4.1.9 does not allow us to define any new total functions. Similar remarks apply to the theory of surjections $\mathbb{N} \twoheadrightarrow \mathbb{N}^{\mathbb{N}}$, considered in C1.2.8. In fact there can never be an epimorphism $N \twoheadrightarrow N^N$ in a non-degenerate topos, by the standard argument used to prove the existence of non-recursive functions: given any morphism $g: N \to N^N$, the composite

$$N \xrightarrow{(g,1)} N^N \times N \xrightarrow{\mathrm{ev}} N \xrightarrow{s} N$$

transposes to yield a morphism $1 \to N^N$ disjoint from the image of g.

As another example of the uses of the canonical signature, we prove the Knaster–Tarski fixed-point theorem and use it to discuss the Cantor–Bernstein theorem (fulfilling a promise made in B2.3.4).

Lemma 4.1.10 (Knaster–Tarski Theorem) *Let A be a complete internal poset in a topos, and suppose $f: A \to A$ is an order-preserving map. Then f has a fixed point; that is, there exists $a: 1 \to A$ such that $fa = a$.*

Proof Consider the subobject $Q \rightarrowtail A$ defined by the term

$$\{x : A \mid \langle x, \ulcorner f \urcorner(x)\rangle \in \ulcorner A_1 \urcorner\} ,$$

where $A_1 \rightarrowtail A \times A$ is the order-relation on A. (Elements of Q are commonly called *pre-fixed points* of f.) Let $a: 1 \to A$ be the composite

$$1 \xrightarrow{\ulcorner Q \urcorner} PA \xrightarrow{\bigvee} A ,$$

where $\bigvee$ is the join map of B2.3.9(ii). Then it is easy to verify successively that fa is an upper bound for Q, that a is a pre-fixed point (i.e. it factors through $Q \rightarrowtail A$), that fa is a pre-fixed point, and finally that a is a fixed point. □

Of course, the fixed point constructed in the proof of 4.1.10 is the unique largest fixed point of f. We may also construct the unique smallest fixed point, by reversing the order relation wherever it appears; and indeed it may be shown that the subobject of fixed points (i.e. the equalizer of f and 1_A) is itself a complete poset in the ordering it inherits from A.

Corollary 4.1.11 (Cantor–Bernstein Theorem) *Suppose given monomorphisms $f: A \rightarrowtail B$ and $g: B \rightarrowtail A$ in a Boolean topos. Then there exists an isomorphism $h: A \to B$.*

Proof Consider the composite

$$PA \xrightarrow{\exists f} PB \xrightarrow{\neg} PB \xrightarrow{\exists g} PA \xrightarrow{\neg} PA\,.$$

This is clearly order-preserving since $\exists f$ and $\exists g$ are order-preserving and $\neg$ is order-reversing, so by 4.1.10 it has a fixed point $a_0 \colon 1 \to PA$. If we write $A_0 \rightarrowtail A$ for the subobject named by a_0, and $A_1 \rightarrowtail A$ for its complement, then it is clear that f maps A_0 isomorphically onto a subobject B_0 of B, and g maps the complement B_1 of B_0 isomorphically onto A_1. So we obtain our required isomorphism h by combining the restriction of f to A_0 with the inverse of the restriction of g to B_1 (and recalling that A is the coproduct of A_0 and A_1). □

The assumption of Booleanness is not necessary for the conclusion of 4.1.11 to hold. For example, in a topos of the form $[\mathcal{C}, \mathbf{Set}_f]$ where $\mathcal{C}$ is a finite category, every object is Dedekind-finite (that is, every monic endomorphism of an object is an isomorphism), and so the Cantor–Bernstein theorem holds vacuously in the sense that the monomorphisms f and g in the statement must already be isomorphisms. On the other hand, the theorem does not hold in all toposes:

Lemma 4.1.12 *Let $\mathcal{E}$ be a topos with a natural number object N. If the Cantor–Bernstein theorem holds in $\mathcal{E}$, then every subobject of an object of $\mathcal{E}$ is an N-indexed union of complemented subobjects.*

Proof Let $A' \rightarrowtail A$ be an arbitrary subobject in $\mathcal{E}$. Consider the objects $B = N \times A$ and $C = B \amalg A'$: we clearly have a monomorphism $\nu_1 \colon B \rightarrowtail C$, but we also have a monomorphism $C \rightarrowtail B$ induced by $s \times 1 \colon B \rightarrowtail B$ and the composite of $A' \rightarrowtail A$ with $o \times 1 \colon A \to B$ (cf. A2.5.8). However, if we have an isomorphism $B \to C$ (or even an epimorphism), then the pullback of $\nu_2 \colon A' \rightarrowtail C$ along it is a complemented subobject of $N \times A$ (that is, an N-indexed family of complemented subobjects of A) whose union is A'. □

Since it is easy to find toposes containing subobjects which are not countable unions of complemented ones (for example, the Sierpiński topos $[\mathbf{2}, \mathbf{Set}]$), we see that the hypothesis of Booleanness cannot be omitted entirely from 4.1.11. In fact B. Banaschewski and G. Brümmer [55] have shown that, if we strengthen the conclusion of the Cantor–Bernstein theorem by requiring that $h_\bullet \leq (f_\bullet \cup g^\bullet)$ in the poset of relations $A \nrightarrow B$ (cf. A3.1.3), then the theorem holds in a topos $\mathcal{E}$ with a natural number object iff $\mathcal{E}$ is Boolean.

Suggestions for further reading: Banaschewski & Brümmer [55], Kock & Reyes [640], Lambek [667], Lambek & Scott [682].

D4.2 λ-Calculus and cartesian closed categories

We now illustrate the general framework which we set up in the previous section by describing the equivalence between theories in the typed λ-calculus and

(small) cartesian closed categories. Historically, the λ-calculus was first introduced (by A. Church [241]) as an untyped 'theory of functions' which it was hoped would provide a foundation for mathematics: the idea was that every term in the calculus should denote a function defined on the whole universe of mathematical discourse, so that it could be applied to any other term (or indeed to itself) to yield a meaningful result. For a detailed account of the untyped λ-calculus and its history, see [67]. We shall encounter it in Chapter F2 (and see also 4.2.15 below), but for the present we shall content ourselves with describing the 'simply typed' λ-calculus, whose aims are much more modest. (Actually, what we describe is not even the simply typed λ-calculus as originally conceived: the latter did not include product-types, but it was soon realized that they could be consistently added – if you have function-types, then the product-type $A \times B$ is implicitly present, because $[A \times B \to C]$ is canonically isomorphic to $[A \to [B \to C]]$ for any C – and from the categorical point of view it is certainly more natural to have the products present from the outset.)

A λ-*signature*, then, is a signature Σ which has the type constructors (i), (ii) and (iii) of Definition 4.1.1; in addition it may have primitive function symbols, but the set Σ-Rel of primitive relation symbols is required to be empty. The λ-*terms* over Σ are those which can be constructed using rules (i)–(vi) of Definition 4.1.2; however, the only formulae which we allow in the λ-calculus are atomic equalities $(s =_A t)$, where s and t are terms of the same type A. Thus the λ-calculus is 'logic-free' in the sense that it contains no logical connectives or quantifiers.

Remark 4.2.1 From another point of view, that of the so-called *Curry–Howard isomorphism* or 'propositions-as-types paradigm', what has happened to the logic in the λ-calculus is that it has been exported from the formulae into the type structure. From this viewpoint, we think of the types over a λ-signature as 'propositions' (the sorts being 'primitive propositions' and the binary operations $\times$ and $[\to]$ corresponding to the logical operations of conjunction and implication), and the function symbols of Σ become the 'axioms'. A λ-term-in-context $\vec{x}.t$ is then thought of as a 'derivation' of the type A of t from the 'hypotheses' $B_1, \dots, B_n$ which are the types of $x_1, \dots, x_n$; and the term constructors of 4.1.2(i)–(vi) correspond to operations on derivations. The equalities which, from the other point of view, are the axioms of our λ-theory, become assertions that two derivations are 'essentially the same'.

A λ-*theory* over Σ is a set of formulae-in-context, i.e. expressions of the form $\vec{x}.(s =_A t)$, where $\vec{x}$ is a context containing all the variables occurring free in s or in t. If we wished to maintain our tradition of axiomatizing theories by means of sequents, we should have to write these in the form $(\top \vdash_{\vec{x}} (s =_A t))$, but since all the sequents we ever consider in the λ-calculus are of this form we might as well omit the $\top$ and the $\vdash$. We shall also, as usual, omit the type subscripts from equality signs whenever we can safely do so; and we shall also feel free to omit

the contexts when (as is almost always the case here) they are the canonical ones for the formulae which follow.

Definition 4.2.2 The rules of inference of the typed λ-calculus are as follows: as usual, the following clauses have the standing hypothesis that the contexts appearing in them contain all the free variables in the terms which follow, and each rule is supposed to apply at all appropriate types.

(*a*) The only *structural rule* (cf. 1.3.1(*a*)) is the substitution rule

$$\frac{\vec{x}\,.\,(s = t)}{\vec{y}\,.\,(s[\vec{r}/\vec{x}] = t[\vec{r}/\vec{x}])}$$

where $\vec{r}$ is any string of terms of the same length and type as $\vec{x}$, and $\vec{y}$ is any context including all the free variables occurring in $\vec{r}$. (Note that, as in 1.3.1(*a*), this rule includes the weakening rule, but not its converse.)

(*b*) The *equality rules* consist of the axiom $x\,.\,(x = x)$, the rules

$$\frac{x, y\,.\,(x = y)}{x, y\,.\,(y = x)} \quad \text{and} \quad \frac{x, y, z\,.\,(x = y) \quad x, y, z\,.\,(y = z)}{x, y, z\,.\,(x = z)}$$

and the rule

$$\frac{\vec{x}\,.\,(s_1 = t_1) \qquad \cdots \qquad \vec{x}\,.\,(s_n = t_n)}{\vec{x}\,.\,(r[\vec{s}/\vec{y}] = r[\vec{t}/\vec{y}])}$$

where $\vec{y}$ is a context containing all the free variables in r. (Note that, since we no longer have finite conjunction in our language, we have had to include the symmetry and transitivity of equality as rules of inference, rather than deducing them as in 1.3.1(*b*).)

(*c*) The *rules for product types* are the axioms $x\,.\,(x =_1 *)$ where x has type 1, and

$$x, y\,.\,(\mathsf{fst}(\langle x, y\rangle) =_A x), \quad x, y\,.\,(\mathsf{snd}(\langle x, y\rangle) =_B y) \quad \text{and}$$
$$z\,.\,(\langle \mathsf{fst}(z), \mathsf{snd}(z)\rangle =_{A\times B} z)$$

where x, y, z have types $A, B, A \times B$ respectively.

(*d*) The *rules for function types* are the two axioms $\vec{x}\,.\,(\mathsf{app}(\lambda y\,.\,s, t) =_B s[t/y])$ (the *β-rule*) and $\vec{x}\,.\,(\lambda y\,.\,\mathsf{app}(t, y) =_{[A\to B]} t)$ where y is a variable of type A not occurring free in t (the *η-rule*), plus the rule

$$\frac{\vec{x}, y\,.\,(s =_B t)}{\vec{x}\,.\,(\lambda y\,.\,s =_{[A\to B]} \lambda y\,.\,t)}$$

which 'strengthens' the third equality rule by eliminating the bound variable $y\colon A$ from the context of its conclusion.

If Σ is a λ-signature, then we may consider Σ-structures (as in 4.1.4) in any cartesian closed category $\mathcal{C}$; and, given such a structure M, we may use clauses (i)–(vii) of Definition 4.1.5 to interpret the λ-terms-in-context over Σ as morphisms of $\mathcal{C}$. Of course, we say an equation-in-context $\vec{x}\,.\,(s = t)$ is *satisfied* in M if the interpretations $[\![\vec{x}\,.\,s]\!]_M$ and $[\![\vec{x}\,.\,t]\!]_M$ are equal, and we say M is a *model* of a λ-theory $\mathbb{T}$ if it satisfies all the axioms of $\mathbb{T}$. We have the usual Soundness Theorem:

Proposition 4.2.3 *If an equation-in-context $\vec{x}\,.\,(s = t)$ is derivable in a λ-theory $\mathbb{T}$, then it is satisfied in all models of $\mathbb{T}$ in cartesian closed categories.*

Proof As usual, we simply have to check that each of the rules of 4.2.2 preserves validity, i.e. that if its hypotheses are all satisfied in some Σ-structure M then so is its conclusion. Most of the axioms were already commented on in 4.1.7; the remaining axioms and rules are equally straightforward. $\square$

To obtain a converse to 4.2.3, we have to construct a syntactic category for an arbitrary λ-theory $\mathbb{T}$. This is actually very much easier than the construction in Section D1.3: because of the logic-free nature of the λ-calculus, there is much less that needs to be constructed. In fact we take the objects of our category $\mathcal{C}_{\mathbb{T}}$ to be simply the types of the signature Σ, and morphisms $A \to B$ are equivalence classes $[x\,.\,t]$ of terms-in-context, where $x\colon A$ and $t\colon B$, and the equivalence relation is $\mathbb{T}$-provable equality (i.e. $[x\,.\,s] = [x\,.\,t]$ iff $x\,.\,(s = t)$ is provable in $\mathbb{T}$). (Of course, we also identify α-equivalent terms-in-context, including the identification $[x\,.\,t] = [y\,.\,t[y/x]]$.) The identity morphism $A \to A$ is simply $[x\,.\,x]$, and composition is given by substitution: the composite of

$$A \xrightarrow{[x\,.\,t]} B \xrightarrow{[y\,.\,s]} C$$

is $[x\,.\,s[t/y]]$. The substitution and equality rules easily imply that this is independent of the choice of the representatives s and t. To verify that it is associative, we need to observe that $r[s/z][t/y]$ is actually the same term as $r[s[t/y]/z]$ provided the variable y does not occur free in r.

Lemma 4.2.4 *The category $\mathcal{C}_{\mathbb{T}}$ just defined is cartesian closed.*

Proof The terminal object of $\mathcal{C}_{\mathbb{T}}$ is the type 1; for any object A, the unique morphism $A \to 1$ is $[x\,.\,*]$. Similarly, the product of A and B is $A \times B$; the product projections are $[z\,.\,\mathsf{fst}(z)]$ and $[z\,.\,\mathsf{snd}(z)]$ (where $z\colon A \times B$), and the morphism $C \to A \times B$ induced by $[w\,.\,s]\colon C \to A$ and $[w\,.\,t]\colon C \to B$ is $[w\,.\,\langle s, t\rangle]$. The exponential B^A is $[A \to B]$; the evaluation map $[A \to B] \times A \to B$ is $[w\,.\,\mathsf{app}(\mathsf{fst}(w), \mathsf{snd}(w))]$, and given any morphism $[z\,.\,t]\colon C \times A \to B$, its transpose $C \to [A \to B]$ is $[w\,.\,\lambda x\,.\,t[\langle w, x\rangle/z]]$. The verification that all of this works is straightforward from the rules in 4.2.2. $\square$

We should perhaps re-emphasize at this point that in the statement of Lemma 4.2.4, and indeed throughout this section, the phrase 'cartesian closed category' means a category having finite products and exponentials, but not necessarily any other finite limits (since the logic-free nature of the λ-calculus means that we do not have any means of constructing such limits syntactically). In the terminology introduced in Section A1.5, we are here dealing with 'improperly' cartesian closed categories, in contrast to what we have been doing almost everywhere else in this book.

Proposition 4.2.5 *The category $\mathcal{C}_{\mathbb{T}}$ contains a Σ-structure $M_{\mathbb{T}}$ which satisfies exactly those equations-in-context that are provable in $\mathbb{T}$. Moreover, for any cartesian closed category $\mathcal{D}$, there is a bijection between natural isomorphism classes of cartesian closed functors $\mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ and isomorphism classes of $\mathbb{T}$-models in $\mathcal{D}$, which is induced in one direction by the mapping $F \mapsto F(M_{\mathbb{T}})$.*

Proof Again, this is all straightforward verification. The structure $M_{\mathbb{T}}$ sends each base sort to itself, and each primitive function symbol $f\colon A \to B$ to $[x \,.\, f(x)]$, where $x\colon A$. An easy induction then shows that for any term-in-context $x \,.\, t$, we have $[\![x \,.\, t]\!]_M = [x \,.\, t]$. (Of course, the language contains terms whose contexts have more than one variable, but we can reduce to the latter by observing that we always have $[\![x, y \,.\, t]\!] = [\![w \,.\, t[\mathsf{fst}(w), \mathsf{snd}(w)/x, y]]\!]$ where w is a variable of the appropriate product type.) Hence the equations-in-context satisfied in $M_{\mathbb{T}}$ are exactly those provable in $\mathbb{T}$. Given a $\mathbb{T}$-model N in an arbitrary cartesian closed category $\mathcal{D}$, the corresponding functor $F_N\colon \mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ sends A to NA, and $[x \,.\, t]$ to $[\![x \,.\, t]\!]_N$; it is clear that F_N is a cartesian closed functor, and that $F_N(M_{\mathbb{T}}) = N$. In the opposite direction, since any cartesian closed functor $F\colon \mathcal{C}_{\mathbb{T}} \to \mathcal{D}$ must preserve interpretations of arbitrary λ-terms-in-context, it is easily seen to be naturally isomorphic to F_N where $N = F(M_{\mathbb{T}})$. □

From 4.2.5, we may deduce various forms of the completeness theorem for λ-theories: for example,

Corollary 4.2.6 *Let $\mathbb{T}$ be a λ-theory, and ϕ an equation-in-context over the signature of $\mathbb{T}$. If ϕ is satisfied in all λ-models in toposes of the form $[\mathcal{C}, \mathbf{Set}]$, then it is provable in $\mathbb{T}$.*

Proof Combine 4.2.5 with the fact (A1.5.6(i)) that the Yoneda embedding $\mathcal{C}_{\mathbb{T}} \to [\mathcal{C}_{\mathbb{T}}^{\mathrm{op}}, \mathbf{Set}]$ is a cartesian closed functor (and faithful, so that it reflects the validity of equations). □

Remark 4.2.7 However, we do not have a 'classical completeness theorem' for λ-theories in terms of models in **Set**, or even in Boolean toposes. Consider the theory over the signature with one base sort A and two function symbols $r\colon A \to [A \to A]$ and $i\colon [A \to A] \to A$, with one axiom asserting that ri is the identity. In any model M of this theory in a Boolean topos, we must have $MA \cong 1$ by 4.1.8(ii), so the equation-in-context $x, y \,.\, (x =_A y)$ will be satisfied.

But in 4.2.16 below, as we remarked after 4.1.8, we shall encounter cartesian closed categories containing objects A such that A^A is a retract of A and $A \to 1$ is not monic; so this equation is not provable in the theory.

In the converse direction to 4.2.5, we have

Example 4.2.8 For any small cartesian closed category $\mathcal{C}$, there is a λ-theory $\mathbb{T}_{\mathcal{C}}$ such that, for any cartesian closed $\mathcal{D}$, isomorphism classes of $\mathbb{T}_{\mathcal{C}}$-models in $\mathcal{D}$ correspond bijectively to natural isomorphism classes of cartesian closed functors $\mathcal{C} \to \mathcal{D}$. To construct $\mathbb{T}_{\mathcal{C}}$, we take a basic sort $\ulcorner A \urcorner$ for each object A of $\mathcal{C}$, and a function symbol $\ulcorner f \urcorner \colon \ulcorner A \urcorner \to \ulcorner B \urcorner$ for each morphism $f \colon A \to B$ in $\mathcal{C}$. (The use of the notation $\ulcorner A \urcorner$ is inescapable here, since we need to distinguish between the types $\ulcorner A \times B \urcorner$ and $\ulcorner A \urcorner \times \ulcorner B \urcorner$, and so on.) In order to ensure that the 'formal' product and function types in Σ coincide with the 'actual' products and exponentials in $\mathcal{C}$, we need a further collection of function symbols, all of which will be denoted by ι (since they 'represent' identity morphisms; strictly speaking, they should be decorated with appropriate subscripts to distinguish them from one another, but in practice they are unlikely to be confused). Explicitly, we have a function symbol $\iota \colon 1 \to \ulcorner 1 \urcorner$ (where the '1' on the right is the terminal object of $\mathcal{C}$, and that on the left is the empty product type); for each pair of objects (A, B), we have $\iota \colon \ulcorner A \urcorner \times \ulcorner B \urcorner \to \ulcorner A \times B \urcorner$, and we also have $\bar{\iota} \colon [\ulcorner A \urcorner \to \ulcorner B \urcorner] \to \ulcorner B^A \urcorner$. The theory $\mathbb{T}_{\mathcal{C}}$ has the following axioms (where we omit the contexts, since they are all canonical):

(*a*) for each object A, $(\ulcorner 1_A \urcorner(x) =_{\ulcorner A \urcorner} x)$, and for each commutative triangle

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd} ,$$

$(\ulcorner g \urcorner(\ulcorner f \urcorner(x)) =_{\ulcorner C \urcorner} \ulcorner h \urcorner(x))$;

(*b*) the axiom $(\iota(*) =_{\ulcorner 1 \urcorner} x)$, and for each pair of objects (A, B), the three axioms

$$(\ulcorner \pi_1 \urcorner(\iota(w)) =_{\ulcorner A \urcorner} \mathsf{fst}(w)), \quad (\ulcorner \pi_2 \urcorner(\iota(w)) =_{\ulcorner B \urcorner} \mathsf{snd}(w)) \quad \text{and}$$
$$(\iota(\langle \ulcorner \pi_1 \urcorner(z), \ulcorner \pi_2 \urcorner(z) \rangle) =_{\ulcorner A \times B \urcorner} z)$$

where w and z have types $\ulcorner A \urcorner \times \ulcorner B \urcorner$ and $\ulcorner A \times B \urcorner$ respectively, and π_1 and π_2 are the product projections in $\mathcal{C}$;

(*c*) for each pair of objects (A, B), the axioms

$$(\ulcorner \mathrm{ev} \urcorner(\iota(\langle \bar{\iota}(w), x \rangle)) =_{\ulcorner B \urcorner} \mathsf{app}(w, x)) \quad \text{and} \quad (\bar{\iota}(\lambda x \,.\, \ulcorner \mathrm{ev} \urcorner(\iota(\langle z, x \rangle))) =_{\ulcorner B^A \urcorner} z)$$

where w, x, z have types $[\ulcorner A \urcorner \to \ulcorner B \urcorner]$, $\ulcorner A \urcorner$ and $\ulcorner B^A \urcorner$ respectively.

It is easy to see that, for any cartesian closed category $\mathcal{D}$, a Σ-structure in $\mathcal{D}$ which satisfies the axioms in (*a*) is (up to isomorphism) 'the same thing as' a functor $F: \mathcal{C} \to \mathcal{D}$ (together with some chosen morphisms $1_{\mathcal{D}} \to F(1_{\mathcal{C}})$, $FA \times FB \to F(A \times B)$ and $FB^{FA} \to F(B^A)$, about which the axioms in (*a*) make no assertions). Moreover, it satisfies the axioms in (*b*) iff the first two types of chosen morphisms are inverse to the canonical comparison maps (equivalently, F preserves finite products); and, given that this happens, it satisfies the axioms in (*c*) iff the chosen morphisms of the third type are inverse to the exponential comparison maps $\theta_{A,B}: F(B^A) \to FB^{FA}$ (equivalently, F is a cartesian closed functor).

Remarks 4.2.9 (*a*) If we start from an arbitrary small cartesian closed category $\mathcal{C}$, construct the theory $\mathbb{T}_{\mathcal{C}}$ as in 4.2.8, and then build the syntactic category of the latter, we will arrive at a category equivalent to $\mathcal{C}$. (It is not isomorphic to $\mathcal{C}$, since it has many more objects; but from every object of the syntactic category we have a canonical isomorphism to one of the form $\ulcorner A \urcorner$ where $A \in \text{ob}\,\mathcal{C}$, namely an appropriate composite of (morphisms denoted by) ι's.) Similarly, if we start from a λ-theory $\mathbb{T}$, build its syntactic category $\mathcal{C}_{\mathbb{T}}$ and then construct a new theory from the latter as in 4.2.8, the new theory will be 'Morita-equivalent' to the old in the sense that there will be natural bijections between the isomorphism classes of their models in arbitrary cartesian closed categories. Thus we have shown that (up to equivalence) *small cartesian closed categories are 'the same thing as' Morita-equivalence classes of λ-theories.*

(*b*) However, we may also construct a λ-theory from something much less than a cartesian closed category. Starting from a directed graph G (as defined before 2.1.1), we may form the signature whose sorts are the vertices of G and whose function symbols are the arrows of G; if we write $\mathbb{T}(G)$ for the empty theory over this signature, then $\mathbb{T}(G)$-models in a cartesian closed category $\mathcal{D}$ are the same thing as graph morphisms $G \to \mathcal{D}$ (cf. 2.1.4(*a*)). Alternatively, if G is actually a (small) category, then we may take the axioms of 4.2.8(*a*) to obtain a theory whose models are functors defined on G. Thus the construction of the syntactic category $\mathcal{C}_{\mathbb{T}}$ in this case yields a left (pseudo-)adjoint for the forgetful functor from (small) cartesian closed categories to small categories. (To make this statement 2-categorically respectable, we must here regard **Cat** and the category of small cartesian closed categories as being made into 'locally grouplike' 2-categories: that is, their only 2-cells are natural isomorphisms. This is because, as we observed after 4.1.4, a non-invertible natural transformation between functors $G \rightrightarrows \mathcal{C}$ will not give rise to a natural transformation between the corresponding cartesian closed functors $\mathcal{C}_{\mathbb{T}(G)} \rightrightarrows \mathcal{D}$.)

An important operation on λ-theories is that of adding a new function symbol to the signature (without adding any new axioms to the theory – though, from the point of view described in 4.2.1, this operation is precisely that of adding a new axiom to the logical system represented by the type structure). Since the effect of adding a function symbol $f: A \to B$ is precisely the same as that

of adding $\overline{f}: 1 \to B^A$, we might as well restrict ourselves to the case when the domain of the function symbol is 1. What effect does this have on the syntactic category of our theory?

Lemma 4.2.10 *Let $\mathbb{T}$ be a λ-theory over a signature Σ, and let $\mathbb{T}[a]$ denote the same theory over the signature obtained by adjoining a new function symbol $a: 1 \to A$ to Σ (where A is a particular type of Σ). Then the syntactic category $\mathcal{C}_{\mathbb{T}[a]}$ is equivalent to the full subcategory $(\mathcal{C}_{\mathbb{T}})_A$ of the slice category $\mathcal{C}_{\mathbb{T}}/A$ whose objects are those of the form A^*B, $B \in \operatorname{ob} \mathcal{C}_{\mathbb{T}}$ (cf.* A1.5.4*).*

Proof Clearly, $\mathcal{C}_{\mathbb{T}[a]}$ has the same objects as $\mathcal{C}_{\mathbb{T}}$; and morphisms $B \to C$ in $\mathcal{C}_{\mathbb{T}[a]}$ are provable-equality-classes of terms-in-context $y \,.\, t$ over the enlarged signature $\Sigma[a]$, where $y: B$ and $t: C$. But we may identify such terms-in-context with terms-in-context $x_0, y \,.\, t'$ over Σ, where x_0 is a particular variable of type A (which is distinct from y even if $A = B$, and which not allowed to become bound), by the simple device of replacing the constant a by the variable x_0 wherever it appears in t. Since the axioms of $\mathbb{T}$ make no mention of a, it is clear that this does not affect the relation of provable equality. Thus we have established a bijection between morphisms $B \to C$ in $\mathcal{C}_{\mathbb{T}[a]}$ and morphisms $A \times B \to C$ in $\mathcal{C}_{\mathbb{T}}$ – equivalently, morphisms $A^*B \to A^*C$ in $\mathcal{C}_{\mathbb{T}}/A$. It is straightforward to verify that this bijection respects composition, and so yields an isomorphism between $\mathcal{C}_{\mathbb{T}[a]}$ and the Kleisli category of the comonad $A \times (-)$ on $\mathcal{C}_{\mathbb{T}}$; but, as we saw in A1.5.4, the latter is equivalent to the full subcategory of $\mathcal{C}/A$ described in the statement. □

In view of 4.2.9(a), we may interpret 4.2.10 as saying that, given a small cartesian closed category $\mathcal{C}$ and an object A of $\mathcal{C}$, the category $\mathcal{C}_A$ of A1.5.4 is the free cartesian-closed-category-with-an-element-of-A generated by $\mathcal{C}$; that is, given any cartesian closed functor $F: \mathcal{C} \to \mathcal{D}$ and any morphism $a: 1 \to FA$ in $\mathcal{D}$, there is a unique extension (up to canonical isomorphism) of F to a cartesian closed functor $\overline{F}: \mathcal{C}_A \to \mathcal{D}$ which sends the diagonal map $A \to A \times A$ (regarded as a morphism $1 \to A^*A$ in $\mathcal{C}_A$) to a. In fact it is not hard to prove this directly, if we identify $\mathcal{C}_A$ with the Kleisli category of $A \times (-)$: we have $\overline{F}(B) = F(B)$ for each object B and, given a morphism $f: B \to C$ in $\mathcal{C}_A$ (that is, a morphism $A \times B \to C$ in $\mathcal{C}$), its image under $\overline{F}$ is the composite

$$FB \cong 1 \times FB \xrightarrow{a \times 1} FA \times FB \xrightarrow{F(f)} FC\,.$$

Of course, it requires verification that this is a (cartesian closed) functor; but that is straightforward.

Lemma 4.2.10 also has an interpretation in terms of the propositions-as-types paradigm of 4.2.1, as a 'functional completeness' result: it says that if you add a new proof a of A to the deduction system represented by (the type structure of) a given λ-theory, then any deduction of C from B in the new system is uniquely

representable as a 'polynomial in a' whose 'coefficients' are deductions in the old system. For more details of this point of view, see [682].

In Section A2.5 we studied the notion of natural number object, and we saw (cf. A2.5.3) that the notion makes sense in any cartesian closed category. It is therefore of interest to consider what happens if we add the type N to the typed λ-calculus. Unfortunately, because of the logic-free nature of the λ-calculus, we cannot fully express the idea that N is a natural number object in it: we can assert the existence of recursively defined morphisms, by means of the term constructor iter_x of 4.1.2(ix), but we cannot assert their uniqueness simply by writing down equations. However, by strengthening the notion of iterator, we can repair at least some of the damage.

Definition 4.2.11 (*a*) By a *λN-signature*, we mean a λ-signature containing the distinguished type N, and having the term-constructors zero and succ defined in 4.1.2(viii), plus a term-constructor $\mathsf{rec}_{x,n}$ (the *recursor*) which operates as follows: given variables $x : A$ (where A is arbitrary) and $n : N$, plus terms $s : A$, $t : A$ and $u : N$ such that neither x nor n is free in t or u, $\mathsf{rec}_{x,n}(s, t, u)$ is a term of type A, in which all occurrences of x and n are bound.

(*b*) The rules of inference of the λN-calculus are those of 4.2.2, plus the axioms $y, \vec{z}\,.\,(\mathsf{rec}_{x,n}(s, y, \mathsf{zero}) = y)$,

$$m, y, \vec{z}\,.\,(\mathsf{rec}_{x,n}(s, y, \mathsf{succ}(m)) = s[\mathsf{rec}_{x,n}(s, y, m), m/x, n])$$

and $m\,.\,(\mathsf{rec}_{x,n}(\mathsf{succ}(x), \mathsf{zero}, m) = m)$. (These correspond to the special case $A = 1$ of the sequents of 4.1.7(iv): in the first two, $\vec{z}$ is a string containing all the free variables of s other than x or n.)

Essentially, the recursor allows the construction of terms defined by recursion as in A2.5.2, and not just as in A2.5.1. The reason why we need this is that the proof of existence of the recursively defined morphism in (the first part of) A2.5.2 used both the existence and uniqueness clauses of A2.5.1. The following proposition is now entirely straightforward, and we omit the details of the proof.

Proposition 4.2.12 *If $\mathbb{T}$ is a theory in the λN-calculus, then the syntactic category $\mathcal{C}_{\mathbb{T}}$ (constructed as in 4.2.4) is cartesian closed and contains a* weak natural number object, *that is an object N equipped with morphisms $o : 1 \to N$ and $s : N \to N$ which satisfies the 'existence' part of the conclusion of* A2.5.2, *but not the 'uniqueness' part.* □

Using the recursor, we may convert the commutative diagrams of A2.5.4 into linguistic definitions of the arithmetic operations on N: for example, we define

addition, the predecessor map and truncated subtraction by

$$\begin{aligned}\mathsf{plus}(x,y) &= \mathsf{rec}_{z,w}(\mathsf{succ}(z),x,y)\ ,\\ \mathsf{pred}(x) &= \mathsf{rec}_{z,w}(w,\mathsf{zero},x)\ ,\ \text{and}\\ \mathsf{tsub}(x,y) &= \mathsf{rec}_{z,w}(\mathsf{pred}(z),x,y)\ .\end{aligned}$$

However, since we lack the uniqueness clause of A2.5.2, we cannot prove that these operations satisfy the usual laws of arithmetic. We may of course add axioms to any λN-theory to assert that particular laws hold; we shall be particularly interested in two of them, namely $x, y\,.\,(\mathsf{plus}(x,y) = \mathsf{plus}(y,x))$ and $x, y\,.\,(\mathsf{tsub}(\mathsf{plus}(x,y),y) = x)$.

The point of these two equations is that they tell us that the type N comes equipped with a *Mal'cev operation*, that is a ternary operation $m(x,y,z) = \mathsf{tsub}(\mathsf{plus}(x,z),y)$ satisfying the equations $m(x,y,y) = x$ and $m(x,x,y) = y$. (Mal'cev operations play an important rôle in universal algebra; see [1114], for example.) In our present context, we shall say that a type A over a λN-signature is a *Mal'cev type* (relative to a particular λN-theory $\mathbb{T}$) if there exists a term m of type A, having three free variables x, y, z of type A, for which the equations-in-context $x, z\,.\,(m[x/y] = z)$ and $x, z\,.\,(m[z/y] = x)$ are derivable in $\mathbb{T}$. Thus, if the two equations at the end of the last paragraph are derivable, N is a Mal'cev type.

The reason for being interested in Mal'cev types is that, if A is such a type, then any term representing a function $N \to A$ has a recursive definition. Specifically, suppose $t = t(n)$ is a term of type A involving a variable n of type N, and let $s = s(x,n)$ be a term of type A with variables x, n of types A, N respectively. (We are departing from our usual custom of not displaying the free variables in a term, in order to improve the readability of the expressions involving multiple substitutions which follow.) By the *Uniqueness Principle* for the pair (s,t), we mean the assertion that, if the equation-in-context $n\,.\,(t(\mathsf{succ}(n)) = s(t(n),n))$ is derivable in our theory, then the equation $n\,.\,(t(n) = \mathsf{rec}_{x,p}(s(x,p), t(\mathsf{zero}), n))$ is derivable.

Now suppose A is a Mal'cev type; let $m(x,y,z)$ be a Mal'cev term for it. Given s and t as above, we define a new term $\tilde{s}(x,n)$ by

$$\tilde{s}(x,n) = m(t(\mathsf{succ}(n)), s(t(n),n), s(x,n))\ .$$

Lemma 4.2.13 *Let A be a Mal'cev type over a λN-signature. Then the Uniqueness Principle holds for all pairs (s,t) iff the equation-in-context*

$$n\,.\,(t(n) = \mathsf{rec}_{x,p}(\tilde{s}(x,p), t(\mathsf{zero}), n))$$

is derivable for all pairs (s,t).

Proof First suppose the Uniqueness Principle holds for the pair $(\tilde{s}, t)$. Since m is a Mal'cev term, we can derive

$$n\,.\,(\tilde{s}(t(n), n) = t(\mathsf{succ}(n)))$$

by substituting $t(n)$ for x in the definition of $\tilde{s}$. So the conclusion of the Uniqueness Principle is derivable; but that is exactly the equation-in-context displayed in the statement of the lemma. Conversely, suppose that this equation is derivable, and that we can also derive $n\,.\,(t(\mathsf{succ}(n)) = s(t(n), n))$. Again using the fact that μ is a Mal'cev term, we derive $x, n\,.\,(\tilde{s}(x, n) = s(x, n))$, so we may substitute s for $\tilde{s}$ in the equation displayed in the statement, to deduce the conclusion of the Uniqueness Principle for (s, t). □

Corollary 4.2.14 *Let $\mathbb{T}$ be a λN-theory in which the equations-in-context $x, y\,.\,(\mathsf{plus}(x, y) = \mathsf{plus}(y, x))$ and $x, y\,.\,(\mathsf{tsub}(\mathsf{plus}(x, y), y) = x)$, and all equations of the form $n\,.\,(t(n) = \mathsf{rec}_{x,p}(\tilde{s}(x, p), t(\mathsf{zero}), n))$ where the type of s and t is a Mal'cev type, are derivable. Then the full subcategory $\mathcal{D}$ of $\mathcal{C}_{\mathbb{T}}$ whose objects are the Mal'cev types relative to $\mathbb{T}$ is a cartesian closed category with a natural number object.*

Proof $\mathcal{D}$ is closed under finite products in $\mathcal{C}_{\mathbb{T}}$, because if $m(x, y, z)$ and $m'(x', y', z')$ are Mal'cev terms for types A and A' respectively, then $\langle m(\mathsf{fst}(x), \mathsf{fst}(y), \mathsf{fst}(z)), m'(\mathsf{snd}(x), \mathsf{snd}(y), \mathsf{snd}(z))\rangle$ is a Mal'cev term for $A \times A'$. Similarly, it is an exponential ideal, since if A has a Mal'cev term $m(x, y, z)$ we may define one for $[B \to A]$, namely $\lambda w\,.\,m(\mathsf{app}(x, w), \mathsf{app}(y, w), \mathsf{app}(z, w))$ where w is of type B. So it is cartesian closed; and, as we have already noted, the two arithmetical equations in the statement ensure that it contains the weak natural number object N of $\mathcal{C}_{\mathbb{T}}$. But then 4.2.13 ensures that N is actually a natural number object in $\mathcal{D}$. □

However, we cannot obtain every (small) cartesian closed category with a natural number object in this way, since it is not true in general that every object in such a category admits a Mal'cev operation.

We conclude this section by considering how small a cartesian closed category $\mathcal{C}$ can be. Clearly, if it has only one object, then that object must be terminal, and the category is not very interesting. Can it have just two objects? If so, then one of them is the terminal object 1, and the other (A, say) must satisfy either $A \times A = A = A^A$ or $A \times A = A$ and $A^A = 1$. The second possibility leads to the conclusion that A is subterminal, and hence that $\mathcal{C}$ is a two-element Heyting algebra, which is again uninteresting; but the first possibility can lead to interesting examples. We note that if $A = A^A$ then 1 is a retract of A (since we always have a morphism $\overline{1_A}\colon 1 \to A^A$), and so we might as well forget about the object 1 for the moment, since we can recover it when necessary by splitting an idempotent (cf. A1.1.8). Thus we are reduced to considering a monoid M (the monoid of endomorphisms of A) equipped with the following structure.

Definition 4.2.15 By a *C-monoid* we mean a monoid M equipped with three distinguished elements p, q and e, a binary operation $\langle -, - \rangle$ and a unary operation $(-)^*$ satisfying the following equations (in which we denote the monoid operation by juxtaposition):

(i) $p\langle x, y\rangle = x$ and $q\langle x, y\rangle = y$ for all x and y.

(ii) $\langle pz, qz\rangle = z$ for all z.

(iii) $e\langle x^*p, q\rangle = x$ and $(e\langle yp, q\rangle)^* = y$ for all x and y.

The idea is that p and q are the product projections $A = A \times A \rightrightarrows A$, $\langle -, - \rangle$ is the binary operation which converts a pair of morphisms $A \rightrightarrows A$ into a single morphism $A \to A \times A = A$, $(-)^*$ converts a morphism $A = A \times A \to A$ into its transpose $A \to A^A = A$, and e is the evaluation map $A = A^A \times A \to A$. Since the theory of C-monoids is algebraic, we may clearly construct nontrivial examples. In particular, given a C-monoid M, we may construct the 'polynomial C-monoid' $M[a]$ freely generated by M together with one extra element a. The functional completeness result 4.2.10 can be used to give an explicit description of the latter: every element of $M[a]$ is of the form $e\langle x(aq)^*, 1\rangle$ for a unique $x \in M$.

Lemma 4.2.16 *Let M be a C-monoid. Then the Cauchy completion $M[\check{\mathcal{E}}]$ of M (cf.* A1.1.9*) is cartesian closed.*

Proof The terminal object is taken to be q^*. To verify that this is indeed idempotent in M, we need the following calculation:

$$\begin{aligned} y^*x &= (e\langle y^*xp, q\rangle)^* \\ &= (e\langle y^*p\langle xp, q\rangle, q\langle xp, q\rangle\rangle)^* \\ &= (e\langle y^*p, q\rangle\langle xp, q\rangle)^* \\ &= (y\langle xp, q\rangle)^* \ , \end{aligned}$$

in which all steps except the third follow directly from the axioms of 4.2.15, and the third is an easy exercise. Substituting q for y, we obtain $q^*x = q^*$, so in particular q^* is idempotent. But this calculation also tells us that q^* is the unique morphism $u \to q^*$ in $M[\check{\mathcal{E}}]$, for any idempotent u.

The product of two idempotents u and v is $\langle up, vq\rangle$, with projections given by up and vq; if $x\colon w \to u$ and $y\colon w \to v$ are morphisms of $M[\check{\mathcal{E}}]$, then the induced morphism $w \to u \times v$ is given by $\langle x, y\rangle$. Similarly, the exponential v^u is defined to be $(ve\langle p, uq\rangle)^*$; the evaluation map $v^u \times u \to v$ is $ve\langle p, uq\rangle$, and the transpose of $x\colon w \times u \to v$ is simply $x^*\colon w \to v^u$. The verification that all of this works is straightforward. □

Of course, M itself appears as the monoid of endomorphisms of the idempotent 1 in $M[\check{\mathcal{E}}]$; this object satisfies $1 = 1 \times 1 = 1^1$ (though it is not the terminal object, unfortunately for the notation!), and so the full subcategory of $M[\check{\mathcal{E}}]$ on this object and the terminal object q^* is also cartesian closed.

C-monoids correspond to the *untyped λ-calculus*, which is defined just as in 4.2.2 except that we omit all the type information (and all references to the

particular type 1 and the corresponding constant $*$). That is, we have a stock of variables, plus a possibly empty set of constants, which we combine into terms using the term constructors fst, snd, $\langle -, - \rangle$, $\lambda x\,.\,-$ and app, and axioms and rules of inference as in 4.2.2, except for the first axiom of 4.2.2(c). (Once again, this is not the untyped λ-calculus as originally conceived by Church, since the latter did not have the term constructors fst, snd or $\langle -, - \rangle$; Lambek and Scott [682] call what we have just described the *extended* λ*-calculus*.) Given a theory $\mathbb{T}$ in this calculus, we may construct its syntactic C-monoid $M_{\mathbb{T}}$, whose elements are the provable-equality-classes of closed terms, with multiplication defined by $[s][t] = [\lambda x\,.\,\mathsf{app}(s, \mathsf{app}(t, x))]$ and identity element $[\lambda x\,.\,x]$. The C-monoid structure is defined by setting $p = [\lambda x\,.\,\mathsf{fst}(x)]$, $q = [\lambda x\,.\,\mathsf{snd}(x)]$, $e = [\lambda x\,.\,\mathsf{app}(\mathsf{fst}(x), \mathsf{snd}(x))]$, and so on. For more information, we refer the reader to [682].

Suggestions for further reading: Barendregt [67], Crole [266], Lambek[668, 673], Lambek & Scott [682], Scott [1104].

D4.3 Toposes as type theories

We are now ready to describe the theories which correspond to (small) toposes in the same way that λ-theories correspond to cartesian closed categories. Normally these would be called 'higher-order intuitionistic type theories', or some similar phrase; however, taking our cue from the name 'λ-theory', we shall simply call them τ*-theories* (the letter τ standing for 'topos' or 'type', according to taste).

A τ*-signature*, then, is one which has the type constructors for finite products and power types, together with the term constructors of 4.1.2(iii), (iv) and (vii). For the moment, we shall not require the signature to have list types, but we shall put them in later as we did with the type N in Section D4.2. Since toposes are cartesian closed, we could if we wished put in function type constructors as well, but since we do not need them we shall keep the logic as simple as possible by leaving them out. In contrast to the λ-calculus, we allow primitive relation symbols as well as function symbols in a τ-signature (though in practice we shall rarely use the former); and we have the full range of (finitary) logical connectives and quantifiers, as specified in 1.1.3(iii)–(x), available in our language, as well as all three types of atomic formulae described in 4.1.3.

Definition 4.3.1 The rules of inference of the τ-calculus are all the first-order rules of 1.3.1(a)–(g), the product type axioms of 4.2.2(c), and the following axioms for power types: $(\top \vdash_w (w = \{x \colon A \mid x \in_A w\}))$ for w a variable of type PA, and $((z \in_A \{y \colon A \mid \phi\}) \dashv\vdash_{\vec{x}, z} \phi[z/y])$ for ϕ any formula with free variables in the string $\vec{x}, y$.

Remark 4.3.2 There is a good deal of redundancy in this presentation of the τ-calculus: in fact we could eliminate all the logical connectives and quantifiers as primitives of our language except for $\Rightarrow$ and $\forall$. To see this let Ω, as usual, denote the type $P1$; if t is a term of type Ω, we shall abbreviate the formula $(* \in_1 t)$

to $\tilde{t}$. (Informally, we think of terms of type Ω as being the same thing as formulae, since the axioms for power types ensure that we have a bijection between provable-equivalence classes of the latter and provable-equality classes of the former; in the other direction, a formula ϕ is mapped to the term $\widehat{\phi} \equiv \{x\colon 1 \mid \phi\}$, where x is a variable not occurring in ϕ.)

We now assert that the sentence $(\forall p : \Omega)\overline{p}$ will do as a substitute for the constant $\bot$: to see this, we have to show that the sequent $(\bot \vdash_{\vec{x}} \phi)$ of 1.3.1(d) is derivable, for an arbitrary ϕ. But we have the following derivation:

$$\dfrac{\dfrac{\dfrac{((\forall p)\overline{p} \vdash_{[]} (\forall p)\overline{p})}{((\forall p)\overline{p} \vdash_{p} \overline{p})}}{((\forall p)\overline{p} \vdash_{\vec{x}} (* \in \{y \mid \phi\}))} \qquad \dfrac{((z \in \{y \mid \phi\}) \vdash_{\vec{x},z} \phi[z/y])}{((* \in \{y \mid \phi\}) \vdash_{\vec{x}} \phi[*/y])}}{((\forall p)\overline{p} \vdash_{\vec{x}} \phi)}$$

where y and z are variables of type 1 not occurring in ϕ. A similar argument shows that $(\forall p\colon \Omega)(\overline{p} \Rightarrow \overline{p})$ will do for $\top$, that $(\phi \vee \psi)$ may be replaced by

$$(\forall p\colon \Omega)((\phi \Rightarrow \overline{p}) \Rightarrow ((\psi \Rightarrow \overline{p}) \Rightarrow \overline{p}))$$

(where p is a variable of type Ω not appearing in ϕ or ψ), and that $(\phi \wedge \psi)$ may similarly be replaced by

$$(\forall p\colon \Omega)((\phi \Rightarrow (\psi \Rightarrow \overline{p})) \Rightarrow \overline{p})\ .$$

(We remark that we have seen the above tricks for deriving finite disjunction from implication and universal quantification over Ω before, in A1.6.4. However, the ones for finite conjunction did not appear there, since we were working in a context where we already knew that finite intersections of subobjects existed.) Finally, we may replace the existential quantification $(\exists x\colon A)\phi$ by

$$(\forall p\colon \Omega)((\forall x\colon A)(\phi \Rightarrow \overline{p}) \Rightarrow \overline{p})\ .$$

This does not exhaust the possible simplifications of 4.3.1. We may also eliminate the term-constructors fst and snd from our language, since a subterm of the form $\mathsf{fst}(t)$, where $t : A \times B$, may be replaced by a new variable $x : A$ satisfying the formula $(\exists y\colon B)(t = \langle x, y\rangle)$. We may even eliminate equality as a primitive symbol, replacing it by the 'Leibnizian' definition that $(s =_A t)$ means $(\forall w : PA)((s \in_A w) \Leftrightarrow (t \in_A w))$, where $\Leftrightarrow$ is defined in the usual way from $\Rightarrow$ and $\wedge$, and w is a variable not occurring in s or t. However, such 'reductionist' tricks run counter to the actual practice of mathematics: a bit of redundancy in one's list of primitives rarely does any harm.

By a τ-*theory* over a signature Σ, we as usual mean a set $\mathbb{T}$ of sequents over Σ, to be thought of as the (non-logical) axioms of the theory. (However, since we have full first-order logic available in the τ-calculus, we could alternatively axiomatize our theories by sets of sentences, replacing a sequent $(\phi \vdash_{\vec{x}} \psi)$ by the

sentence $(\forall \vec{x})(\phi \Rightarrow \psi)$. In practice, if a sequent of the form $(\top \vdash_{[]} \phi)$ is derivable in a τ-theory $\mathbb{T}$, we shall tend to say simply that the sentence ϕ is derivable in $\mathbb{T}$.) And we have the usual notion of a *model* for a τ-theory $\mathbb{T}$ in a topos $\mathcal{E}$, namely a structure M for its signature (defined as in 4.1.4) which satisfies all the axioms of $\mathbb{T}$. The Soundness Theorem is yet again straightforward:

Proposition 4.3.3 *Let $\mathbb{T}$ be a τ-theory over a signature Σ. If a sequent σ is derivable in $\mathbb{T}$ (using the rules specified in* 4.3.1*), then it is satisfied in any model of $\mathbb{T}$ in a topos.*

Proof As usual, we have to verify soundness of each of the individual axioms and rules of inference. But the soundness of the first-order axioms (in arbitrary Heyting categories) was verified in 1.3.2, and that of the axioms for product and power types in 4.1.7(i) and (iii). □

To obtain a converse to 4.3.3, we seek as usual to construct a syntactic category $\mathcal{E}_{\mathbb{T}}$ from an arbitrary τ-theory $\mathbb{T}$, and to prove that it is a topos (and that it contains a conservative model of $\mathbb{T}$). With the extra logical structure now at our disposal, we may somewhat simplify the construction of this category, as compared with the first-order syntactic categories which we constructed in Section D1.4. In the first place, we note that, for the objects of our category, it suffices to consider (α-equivalence classes of) formulae-in-context $x\,.\,\phi$ where the context is a single variable x; for, in any Σ-structure, the interpretation of $x, y\,.\,\phi$ will coincide with (or at least be canonically isomorphic to) that of $z\,.\,\phi[\mathsf{fst}(z), \mathsf{snd}(z)/x, y]$ where z is a variable whose type is the product of those of x and y. (Similarly, a sentence ϕ in the empty context may be replaced by $x\,.\,\phi$, where x is a variable of type 1.) But then we may also exploit the correspondence between (provable-equivalence classes of) formulae in a context with one variable x of type A, and (provable-equality classes of) closed terms of type PA (which sends $x\,.\,\phi$ to $\{x : A \mid \phi\}$ and a closed term a to $x\,.\,(x \in_A a)$), and thus regard the objects of our syntactic category as α-equivalence classes of closed terms whose type is a power type PA for some A. In this way, we may eliminate the contexts altogether from our description of $\mathcal{E}_{\mathbb{T}}$. (In truth, this is only a very minor notational change: the object of the syntactic category which was denoted $\{x\,.\,\phi\}$ in Section D1.4 will now be denoted $\{x \mid \phi\}$.) It will be convenient to have a notation for (the object of $\mathcal{E}_{\mathbb{T}}$ corresponding to) the particular closed term $\{x : A \mid \top\}$, where A is a type of Σ; we shall denote it by $\overline{A}$.

If a and b are closed terms of types PA and PB respectively, then a morphism $a \to b$ in $\mathcal{E}_{\mathbb{T}}$ will be represented by a closed term f of type $P(A \times B)$, satisfying the usual functionality axioms (suitably translated into our new terminology): in other words, such that the sequents

$$(((\langle x, y \rangle \in f) \vdash_{x,y} ((x \in a) \wedge (y \in b))) ,$$
$$((x \in a) \vdash_x (\exists y)((\langle x, y \rangle \in f)) , \text{ and}$$
$$((((\langle x, y \rangle \in f) \wedge ((\langle x, y' \rangle \in f)) \vdash_{x,y,y'} (y = y'))$$

are provable in $\mathbb{T}$. Two such terms f and f' will of course represent the same morphism iff the sentence $(f = f')$ is provable in $\mathbb{T}$; as usual, we shall write $[f]$ for the provable-equality class of a closed term f. The composite of two morphisms $[f]: a \to b$ and $[g]: b \to c$ is the provable-equality class of the closed term

$$\{z\colon A \times C \mid (\exists y\colon B)((\langle \mathsf{fst}(z), y\rangle \in f) \wedge (\langle y, \mathsf{snd}(z)\rangle \in g))\} .$$

Just as in the first-order case, it is straightforward to verify that this composition is well-defined and associative, and that the identity morphism on an object a is given by $[\{z\colon A \times A \mid ((\mathsf{fst}(z) = \mathsf{snd}(z)) \wedge (\mathsf{fst}(z) \in a))\}]$; so we have

Lemma 4.3.4 *$\mathcal{E}_{\mathbb{T}}$ is a category.* □

It is similarly straightforward to establish

Lemma 4.3.5 *$\mathcal{E}_{\mathbb{T}}$ has finite limits.*

Proof The terminal object may be taken to be $\overline{1}$ (or $\{x\colon 1 \mid (x = *)\}$, if you prefer); given an arbitrary object $a\colon PA$, it is easy to see that a term $f\colon P(A \times 1)$ represents a morphism $a \to \overline{1}$ iff it is provably equal to $\{z \mid \mathsf{fst}(z) \in a\}$. We define the product of two objects $a : PA$ and $b : PB$ to be the closed term $\{z\colon A \times B \mid ((\mathsf{fst}(z) \in a) \wedge (\mathsf{snd}(z) \in b))\}$; the first product projection $a \times b \to a$ is represented by the term

$$\{w\colon (A \times B) \times A \mid ((\mathsf{fst}(w) \in a \times b) \wedge (\mathsf{fst}(\mathsf{fst}(w)) = \mathsf{snd}(w)))\} ,$$

with a similar expression for the second projection. Given morphisms $[f]: c \to a$ and $[g]: c \to b$, where $c : PC$, the induced morphism $([f], [g]): c \to a \times b$ is represented by the term

$$\{w\colon C \times (A \times B) \mid ((\langle \mathsf{fst}(w), \mathsf{fst}(\mathsf{snd}(w))\rangle \in f) \wedge (\langle \mathsf{fst}(w), \mathsf{snd}(\mathsf{snd}(w))\rangle \in g))\} .$$

Given two morphisms $[f], [g]: a \rightrightarrows b$, the domain ($e$, say) of their equalizer is the object

$$\{x\colon A \mid (\exists y\colon B)((\langle x, y\rangle \in f) \wedge (\langle x, y\rangle \in g))\} ,$$

and the monomorphism $e \rightarrowtail a$ is represented by $\{z\colon A \times A \mid ((\mathsf{fst}(z) = \mathsf{snd}(z)) \wedge (\mathsf{fst}(z) \in e))\}$. (As in 1.4.3, we note that the construction of equalizers depends on the choice of representatives f and g, not just on their provable-equality classes; but we could eliminate these choices by defining objects as well as morphisms to be provable-equality classes of terms.) As usual, the verification that all of this works is tedious but straightforward. □

Lemma 4.3.6 *Let $[f]: a \to b$ be a morphism of $\mathcal{E}_{\mathbb{T}}$.*

(i) *$[f]$ is an isomorphism iff the sequents*

$$((y \in b) \vdash_y (\exists x)(\langle x, y\rangle \in f)) \quad \textit{and}$$
$$(((\langle x, y\rangle \in f) \wedge (\langle x', y\rangle \in f)) \vdash_{x,x',y} (x = x'))$$

are provable in $\mathbb{T}$.

(ii) *$[f]$ is a monomorphism iff the second sequent above is provable in $\mathbb{T}$.*

Proof (i) If the sequents above are provable, then the term $f^\circ = \{z : B \times A \mid (\langle \mathsf{snd}(z), \mathsf{fst}(z)\rangle \in f)\}$ is easily seen to represent a morphism $b \to a$, which is inverse to f. Conversely, if $[f][g]$ and $[g][f]$ are both identity morphisms in $\mathcal{E}_{\mathbb{T}}$ then g is provably equal to f°, and so f must satisfy the sequents in the statement.

(ii) Following the constructions of 4.3.5, we see that (the domain of) the kernel-pair of $[f]$ is $\{z : A \times A \mid (\exists y : B)(((\langle \mathsf{fst}(z), y\rangle \in f) \wedge (\langle \mathsf{snd}(z), y\rangle \in f))\}$, and the second sequent is equivalent to saying that the terms representing the two projections from this object to a are provably equal. □

Unsurprisingly, it is also possible to show that a morphism $[f]$ is epic in $\mathcal{E}_{\mathbb{T}}$ iff the first of the two sequents of 4.3.6(i) is provable; but we shall not need this.

Corollary 4.3.7 *Let $a : PA$ be an object of $\mathcal{E}_{\mathbb{T}}$. Then any subobject of a is isomorphic to one of the form $[i]: a' \rightarrowtail a$, where $a' : PA$ is a term such that $((x \in a') \vdash_x (x \in a))$ is provable in $\mathbb{T}$ and i is the term*

$$\{z : A \times A \mid ((\mathsf{fst}(z) = \mathsf{snd}(z)) \wedge (\mathsf{fst}(z) \in a'))\} .$$

Moreover, for two such subobjects $[i]: a' \rightarrowtail a$ and $[i']: a'' \rightarrowtail a$, we have $a' \leq a''$ in Sub(A) *iff the sequent $((x \in a') \vdash_x (x \in a''))$ is provable in $\mathbb{T}$.*

Proof First, we note from 4.3.6(ii) that any morphism of the specified form is monic. Given an arbitrary monomorphism $[f]: b \rightarrowtail a$, let a' be the term $\{x : A \mid (\exists y : B)(\langle y, x\rangle \in f)\}$. Then it is straightforward to verify that f also represents a morphism $b \to a'$ which is a factorization of $[f]$ through $[i]$, and that this factorization is an isomorphism. The final assertion is similarly straightforward, since any factorization of $[i]$ through $[i']$ must be represented by the term i. □

Proposition 4.3.8 *$\mathcal{E}_{\mathbb{T}}$ is a topos.*

Proof After 4.3.5, it suffices (by A2.3.4) to show that $\mathcal{E}_{\mathbb{T}}$ has power objects. Given an object $a : PA$, we define its power object Pa to be the term $\{w : PA \mid (\forall x : A)((x \in w) \Rightarrow (x \in a))\}$, of type PPA. The membership relation $\in_a$ is defined to be $\{z : PA \times A \mid ((\mathsf{fst}(z) \in Pa) \wedge (\mathsf{snd}(z) \in \mathsf{fst}(z)))\}$, made into a subobject of $Pa \times a$ as in 4.3.7. Given an arbitrary subobject $c \rightarrowtail b \times a$, which we may assume to be represented as in 4.3.7 (so that c is a closed term of type

$P(B \times A))$, we define its name $b \to Pa$ to be the morphism represented by the term

$$\{z\colon B \times PA \mid ((\mathsf{fst}(z) \in b) \wedge (\mathsf{snd}(z) = \{x\colon A \mid (\langle \mathsf{fst}(z), x\rangle \in c)\}))\} \,.$$

It is straightforward to verify that this does represent a morphism, and that the latter is the unique morphism $b \to Pa$ for which there is a pullback square

$$\begin{tikzcd} c \arrow[r] \arrow[d, tail] & \in_a \arrow[d, tail] \\ b \times a \arrow[r] & Pa \times a \end{tikzcd}$$

in $\mathcal{E}_{\mathbb{T}}$. □

Remark 4.3.9 As we observed for the first-order case in 1.4.15, we could have been still more economical in our construction of $\mathcal{E}_{\mathbb{T}}$, by exploiting the theory of allegories developed in Chapter A3 – that is, by constructing it as the category of maps of a suitable power allegory. For this approach, we begin by defining an allegory $\mathcal{A}_{\mathbb{T}}$, whose objects are simply the types of Σ, and whose morphisms $A \looparrowright B$ are provable-equality classes of closed terms r of type $P(A \times B)$, ordered by $[r] \leq [s]$ iff $((z \in r) \vdash_z (z \in s))$ is provable in $\mathbb{T}$. The composite of two morphisms $[r]\colon A \looparrowright B$ and $[s]\colon B \looparrowright C$ is the provable-equality class of

$$\{z\colon A \times C \mid (\exists y\colon B)((\langle \mathsf{fst}(z), y\rangle \in r) \wedge (\langle y, \mathsf{snd}(z)\rangle \in s))\} \,,$$

and the opposite of $[r]$ is the class of

$$\{z\colon B \times A \mid (\langle \mathsf{snd}(z), \mathsf{fst}(z)\rangle \in r)\} \,.$$

It is straightforward to verify that $\mathcal{A}_{\mathbb{T}}$ is an allegory, and in fact a division allegory in the sense of A3.4.1: given $[r]\colon A \looparrowright B$ and $[s]\colon A \looparrowright C$, we define $[s]/[r]\colon B \looparrowright C$ to be the provable-equality class of

$$\{z\colon B \times C \mid (\forall x\colon A)((\langle x, \mathsf{fst}(z)\rangle \in r) \Rightarrow (\langle x, \mathsf{snd}(z)\rangle \in s))\} \,.$$

Indeed, $\mathcal{A}_{\mathbb{T}}$ is a power allegory in the sense of A3.4.5: the power object of an object A is none other than the power type PA, and the canonical morphism $\in_A\colon PA \looparrowright A$ is the provable-equality class of $\{z \mid (\mathsf{snd}(z) \in \mathsf{fst}(z))\}$. $\mathcal{A}_{\mathbb{T}}$ is not in general tabular; but it is pre-tabular, since the hom-poset $\mathcal{A}_{\mathbb{T}}(A, B)$ has a top element $[\{z : A \times B \mid \top\}]$, which is tabulated by the product projections $A \times B \to A$ and $A \times B \to B$ (we leave to the reader the task of defining these and verifying that they are maps). Its hom-posets also have bottom elements (the bottom element of $\mathcal{A}_{\mathbb{T}}(A, B)$ being $[\{z : A \times B \mid \bot\}]$), so by A3.4.9 we obtain a topos by splitting the set $\mathcal{K}$ of cores in $\mathcal{A}_{\mathbb{T}}$ and then cutting down to

the category of maps of $\mathcal{A}_{\mathbb{T}}[\check{\mathcal{K}}]$. And it is not hard to verify that, if we carry through the above constructions, we arrive at a category equivalent to $\mathcal{E}_{\mathbb{T}}$ as defined earlier in this section.

Next, we define the canonical Σ-structure $M_{\mathbb{T}}$ in $\mathcal{E}_{\mathbb{T}}$. For a base sort A, we define $M_{\mathbb{T}}A$ to be $\overline{A}$. If A is a compound type, it is not true that $M_{\mathbb{T}}A$ is simply $\overline{A}$, but an easy induction shows that it is always of the form $\{x\colon A \mid \phi\}$ for some formula ϕ such that $(\top \vdash_x \phi)$ is provable (in the empty theory over Σ), and hence it is canonically isomorphic to $\overline{A}$. If $f\colon A \to B$ is a function symbol of Σ, we define $M_{\mathbb{T}}f$ to be $[\{z\colon A \times B \mid (f(\mathsf{fst}(z)) = \mathsf{snd}(z))\}]$ (which is easily checked to be a morphism $MA \to MB$ in $\mathcal{E}_{\mathbb{T}}$); similarly, if $R \rightarrowtail A$ is a relation symbol, $M_{\mathbb{T}}R$ is $\{x\colon A \mid R(x)\}$, made into a subobject of MA as in 4.3.7.

Lemma 4.3.10

(i) *If $x\,.\,t$ is any term-in-context over Σ (where the context has just one variable $x : A$, and $t : B$, say), then $[\![x\,.\,t]\!]_M$ is the morphism $MA \to MB$ represented by the term $\{z\colon A \times B \mid (\mathsf{snd}(z) = t[\mathsf{fst}(z)/x])\}$.*

(ii) *If $x\,.\,\phi$ is any formula-in-context over Σ (again, having a context with just one variable $x\colon A$), then $[\![x\,.\,\phi]\!]_M$ is the subobject of MA corresponding (as in 4.3.7) to the term $\{x\colon A \mid \phi\}$ of type PA.*

(iii) *A sequent $(\phi \vdash_{\vec{x}} \psi)$ over Σ is satisfied in $M_{\mathbb{T}}$ iff it is provable in $\mathbb{T}$.*

Proof (i) and (ii) are proved by a straightforward (simultaneous) induction. For (iii), we note that the axioms for product types enable us to replace any sequent $(\phi \vdash_{\vec{x}} \psi)$ by a sequent $(\phi' \vdash_z \psi')$ whose context has a single variable, such that the former is provable in $\mathbb{T}$ iff the latter is. The result is then immediate from (ii) and the last assertion of 4.3.7. □

From (iii), we immediately deduce that M is a conservative model of $\mathbb{T}$. Hence we have

Corollary 4.3.11 (Completeness Theorem for τ-calculus) *Let $\mathbb{T}$ be a τ-theory, and σ a sequent over the signature of $\mathbb{T}$. If σ is satisfied in all models of $\mathbb{T}$ in toposes, then it is provable in $\mathbb{T}$.* □

Remark 4.3.12 Given a (finitary) first-order theory $\mathbb{T}$, we may of course regard it as a τ-theory whose axioms all happen to lie in the first-order part of the language over its signature. Thus 4.3.11 almost contains an alternative proof of the result, already noted in 3.1.18, that validity in models in toposes is sufficient to determine provability in $\mathbb{T}$. ('Almost' because the notion of provability has changed, in that we have added the axioms for power types: of course, we should not expect these to allow us to derive any new first-order conclusions from first-order axioms, but it is not *a priori* obvious that they do not.)

Given an arbitrary topos $\mathcal{F}$ and a τ-theory $\mathbb{T}$, we write $\mathbb{T}\text{-Mod}(\mathcal{F})$ for the groupoid of $\mathbb{T}$-models in $\mathcal{F}$ and isomorphisms between them. (The reason for

restricting to isomorphisms was explained after 4.1.4.) And we write $\mathfrak{Log}$ for the 2-category of toposes, logical functors and natural isomorphisms between them. Any logical functor $\mathcal{F} \to \mathcal{G}$ induces a functor $\mathbb{T}\text{-Mod}(\mathcal{F}) \to \mathbb{T}\text{-Mod}(\mathcal{G})$; this is not quite trivial, because in our definition of a Σ-structure M we required that we should have equalities (not just isomorphisms) $M(A \times B) = MA \times MB$ and $M(PA) = P(MA)$, and a logical functor need not strictly preserve these. However, since our collection of types was built up recursively from the base sorts, we may define FM by setting $FM(A)$ equal to $F(MA)$ for each base sort A, and then for each compound type we have a canonical isomorphism from $F(MA)$ to $FM(A)$, across which we may transport the rest of the structure. Now we have

Proposition 4.3.13 *For any topos $\mathcal{F}$ and any τ-theory $\mathbb{T}$, the assignment $F \mapsto F(M_{\mathbb{T}})$ is part of an equivalence of categories*

$$\mathfrak{Log}\,(\mathcal{E}_{\mathbb{T}}, \mathcal{F}) \simeq \mathbb{T}\text{-Mod}(\mathcal{F})\ .$$

Proof We construct a functor in the opposite direction, as follows. Given a $\mathbb{T}$-model N in $\mathcal{F}$, we define a functor $F_N : \mathcal{E}_{\mathbb{T}} \to \mathcal{F}$ by sending an object a of $\mathcal{E}_{\mathbb{T}}$ to the subobject of NA named by $[\![a]\!]_N : 1 \to N(PA) = P(NA)$, and a morphism $[f]: a \to b$ to the morphism $F_N(a) \to F_N(b)$ whose graph is the object named by $[\![f]\!]_N$. (The soundness theorem 4.3.3 ensures that the latter is indeed the graph of a morphism $F_N(a) \to F_N(b)$, and that (up to isomorphism) it is independent of the choice of the representative f.) The verification that F_N is a logical functor is immediate from the definition of the topos structure on $\mathcal{E}_{\mathbb{T}}$ in 4.3.5 and 4.3.8 (and the soundness theorem, again), and the verification that the assignment $N \mapsto F_N$ is functorial is equally straightforward. Since any logical functor preserves the interpretations of arbitrary terms and formulae over Σ, it is clear that if F is any such functor $\mathcal{E}_{\mathbb{T}} \to \mathcal{F}$ then we have isomorphisms $F(a) \cong F_{F(M)}(a)$ for all objects a of $\mathcal{E}_{\mathbb{T}}$, which form a natural isomorphism $F \cong F_N$ where $N = F(M_{\mathbb{T}})$. Similarly, if N is any $\mathbb{T}$-model in $\mathcal{F}$, we have a canonical isomorphism $N \cong F_N(M_{\mathbb{T}})$. □

Examples 4.3.14 (*a*) Proposition 4.3.13 can be thought of as saying that $\mathcal{E}_{\mathbb{T}}$ is the topos 'freely generated' by the $\mathbb{T}$-model $M_{\mathbb{T}}$. Even in the case when $\mathbb{T}$ is the empty theory over the 'empty signature', this is a nontrivial assertion: it says that the 2-category $\mathfrak{Log}$ has an initial object (in the 2-categorical sense). (The 'empty signature' is not literally empty: although it has no base sorts (and no primitive function symbols, or primitive relations apart from membership and equality), it has many types such as 1, $P1$, $PP1$, $P1 \times P1$ and so on; thus its collection of terms and formulae is already quite a rich one.) The initial object of $\mathfrak{Log}$ is often called simply the *free topos*; we shall devote much of Chapter F3 to further study of it (and its close relatives).

(*b*) Slightly more generally, if G is any small directed graph, we may construct the free topos generated by G, in the sense that logical functors from it to a topos $\mathcal{F}$ correspond to diagrams of shape G in $\mathcal{F}$. To do this, we simply take the

empty τ-theory over the signature whose base sorts are the vertices of G, with the arrows of G taken as function symbols.

(c) More generally still, given any finitary sketch $\mathbb{S}$ (in the sense of 2.1.2(d)), we may construct a topos $\mathcal{E}_{\mathbb{S}}$ such that logical functors $\mathcal{E}_{\mathbb{S}} \to \mathcal{F}$ correspond to $\mathbb{S}$-models in $\mathcal{F}$. (This is not quite as dramatic as it seems, because we lose most of the categorical structure of $\mathbb{S}$-Mod($\mathcal{F}$): we can 'see' only the isomorphisms in this category as natural isomorphisms between logical functors. However, following on from 2.4.12(c), we note that there exists a topos $\mathcal{E}$ for which there are logical functors $\mathcal{E} \to \mathbf{Set}$ with arbitrarily large finite groups of natural automorphisms, but none with infinite automorphism groups; thus the theory of logical functors defined on $\mathcal{E}$ cannot be Morita-equivalent to a finitary first-order theory.) It is easy to see how we may write down the conditions that certain finite diagrams commute, or that certain finite cones are limits, as sequents over the signature described in (b) (enriched with suitable additional function symbols ι as in 4.2.9; cf. 4.3.15 below). To say that certain finite cocones are colimits, we exploit the monadicity of the contravariant power-object functor $P\colon \mathcal{F}^{\mathrm{op}} \to \mathcal{F}$ (A2.2.7), which implies that a (finite) diagram in $\mathcal{F}$ is a colimit iff its image under P is a limit.

(d) Suppose given a τ-theory $\mathbb{T}$ over a signature Σ, and let A be a particular type of Σ. As with λ-theories (cf. 4.2.11), we often wish to consider the theory $\mathbb{T}[a]$ obtained by adding a new function symbol $a\colon 1 \to A$ to Σ, and no new axioms. Thanks to the 'fundamental theorem of topos theory' (A2.3.2), we may identify the syntactic category of $\mathbb{T}[a]$ (up to equivalence) without doing any hard work: *$\mathcal{E}_{\mathbb{T}[a]}$ is equivalent to the slice category $\mathcal{E}_{\mathbb{T}}/\overline{A}$.* For, given a logical functor $F\colon \mathcal{E} \to \mathcal{F}$ and a morphism $b\colon 1 \to F(B)$ in $\mathcal{F}$, we obtain a logical functor

$$\mathcal{E}/B \xrightarrow{F/B} \mathcal{F}/F(B) \xrightarrow{b^*} \mathcal{F}\ ;$$

and conversely, given $G\colon \mathcal{E}/B \to \mathcal{F}$, we have the composite

$$\mathcal{E} \xrightarrow{B^*} \mathcal{E}/B \xrightarrow{G} \mathcal{F}$$

and the morphism $G(\Delta)\colon 1 \to G(B^*B)$, where Δ is the diagonal map of B regarded as a morphism $1_B \to B^*B$ in $\mathcal{E}/B$. It is easy to see that these two constructions are inverse to each other up to natural isomorphism; so, applying the result when $\mathcal{E} = \mathcal{E}_{\mathbb{T}}$ and $B = \overline{A} \cong M_{\mathbb{T}}A$, we deduce that logical functors defined on $\mathcal{E}_{\mathbb{T}}/\overline{A}$ correspond to models of $\mathbb{T}[a]$.

(e) A very similar argument shows that, if $\mathbb{T}'$ is obtained from $\mathbb{T}$ by adding a single sentence ϕ as an additional axiom (and leaving the signature unchanged), then $\mathcal{E}_{\mathbb{T}'}$ may be identified with $\mathcal{E}_{\mathbb{T}}/[\![\phi]\!]_M$. In particular, we note that the free topos of (a) contains a subterminal object (the interpretation of the sentence $(\forall p : \Omega)(\overline{p} \vee \neg\overline{p})$), such that the corresponding slice category is initial (in the 2-categorical sense) in the 2-category $\mathfrak{BooLog}$ of Boolean toposes and logical functors. In fact the initial object of this 2-category is not hard to identify

explicitly; it is equivalent to $\mathbf{Set}_f$, since if $\mathcal{B}$ is a Boolean topos then the unique coherent functor $\mathbf{Set}_f \to \mathcal{B}$ which we constructed in A1.4.7 is actually logical. (However, things become more interesting once we add a natural number object; cf. 4.3.19(d) below.)

Proposition 4.3.13 is one half of the assertion that τ-theories are 'the same thing as' small toposes, in the same sense that λ-theories are the same thing as small cartesian closed categories. For the other half, we need

Proposition 4.3.15 *Given a small topos $\mathcal{E}$, there exists a τ-theory $\mathbb{T}_{\mathcal{E}}$ such that for any topos $\mathcal{F}$ we have*

$$\mathfrak{Log}\,(\mathcal{E},\mathcal{F}) \simeq \mathbb{T}_{\mathcal{E}}\text{-Mod}(\mathcal{F}) \ ,$$

naturally in $\mathcal{F}$.

Proof The construction is very similar to that of 4.2.9, and we shall not give it in great detail. For our signature, we take (i) a basic sort $\ulcorner A \urcorner$ for each object A of $\mathcal{E}$; (ii) a function symbol $\ulcorner f \urcorner \colon \ulcorner A \urcorner \to \ulcorner B \urcorner$ for each morphism $f \colon A \to B$ in $\mathcal{E}$; (iii) a function symbol $\iota \colon 1 \to \ulcorner 1 \urcorner$ and function symbols $\iota \colon \ulcorner A \urcorner \times \ulcorner B \urcorner \to \ulcorner A \times B \urcorner$ for each pair of objects (A, B); and (iv) a function symbol $\tilde{\iota} \colon P\ulcorner A \urcorner \to \ulcorner PA \urcorner$ for each object A. The axioms are (the sequents corresponding to) the equations taken as axioms in 4.2.9(a) and (b), plus the following:

(c) for each monomorphism $m \colon A' \rightarrowtail A$ in $\mathcal{E}$ with classifying map $\phi \colon A \to \Omega$, the two axioms $((\ulcorner m \urcorner(y) = \ulcorner m \urcorner(y')) \vdash_{y,y'} (y = y'))$ and

$$((\ulcorner \phi \urcorner(x) = \ulcorner \top \urcorner(\iota(*))) \vdash_x (\exists y)(\ulcorner m \urcorner(y) = x));$$

(d) for each object A, the axioms

$$((x \in w) \dashv\vdash_{x,w} (\ulcorner \mathrm{ev} \urcorner(\iota(\langle \tilde{\iota}(w), x \rangle)) = \ulcorner \top \urcorner(\iota(*))))$$

where x and w are variables of types $\ulcorner A \urcorner$ and $P\ulcorner A \urcorner$ respectively, and

$$(\top \vdash_u (u = \tilde{\iota}(\{x \colon A \mid (\ulcorner \mathrm{ev} \urcorner(\iota(\langle u, x \rangle)) = \ulcorner \top \urcorner(\iota(*)))\})))$$

where u is of type $\ulcorner PA \urcorner$. (Here ev denotes the evaluation map $PA \times A \to \Omega$ in $\mathcal{E}$.)

Now we saw in 4.2.9 that a model M (in a topos $\mathcal{F}$, say) for the axioms in groups (a) and (b) is essentially the same thing as a functor $F \colon \mathcal{E} \to \mathcal{F}$ preserving finite products; such a model satisfies the axioms in (c) iff the corresponding functor F preserves arbitrary pullbacks of $\top \colon 1 \rightarrowtail \Omega$ (which, by a remark that we made following A1.6.1, is equivalent to saying that it preserves all finite limits); and finally M satisfies the axioms in (d) iff the interpretation in M of $\tilde{\iota}$ yields a two-sided inverse for the canonical comparison map $\phi_A \colon F(PA) \to P(FA)$ for each A, i.e. iff F is a logical functor. □

As we did with λ-theories in 4.2.13, we now wish to extend our notion of τ-theory by adding list types (or at least the particular list type $N = L1$) to our signature. (Since, as we saw in A2.5.17, any topos with a natural number object has list objects – and since any logical functor between toposes with natural number objects preserves the natural number object by A2.2.10(i) and A2.5.6(ii), and hence also preserves list objects – it does not make any essential difference which we do.)

Definition 4.3.16 (*a*) By a *τL-signature*, we mean a higher-order signature Σ having the type constructors of 4.1.1(i), (ii), (iv) and (v), and the term constructors of 4.1.2(i)–(iv) and (vii)–(ix), plus the three types of atomic formulae of 4.1.3 and the full range of (finitary) logical connectives and quantifiers. A *τL-theory* is a set of sequents (or sentences, if you prefer) over such a signature.

(*b*) The rules of the *τL-calculus* are those of the τ-calculus as given in 4.3.1, plus the following rules for list types: the axioms

$$(\top \vdash_{\vec{z},w} (\mathsf{iter}_{x,y}(s, w, [\,]) = w)) \quad \text{and}$$
$$(\top \vdash_{\vec{z},u,v,w} (\mathsf{iter}_{x,y}(s, w, \mathsf{cons}(u, v)) = s[\mathsf{iter}_{x,y}(s, w, v), u/x, y])) \ ,$$

where $\vec{z}$ is any context containing the free variables of s other than x and y, and 'Peano's fifth postulate'

$$((([\,] \in z) \wedge (\forall x)(\forall y)((y \in z) \Rightarrow (\mathsf{cons}(x, y) \in z))) \vdash_z (\forall y)(y \in z))$$

where x, y, z are of types $A, LA, P(LA)$ respectively.

If we are merely given the particular list type N in our signature, we shall call it a τN-signature, and speak of τN-theories and the τN-calculus. In practice, we usually work with the latter; but for many purposes it is no harder to work with arbitrary list types.

Remarks 4.3.17 (*a*) The Soundness Theorem for the τL-calculus is as usual easy to prove: *if σ is provable in a τL-theory $\mathbb{T}$, then it is satisfied in all models of $\mathbb{T}$ in toposes with natural number objects.* For we verified the soundness of the first two axioms of 4.3.16(*b*) in 4.1.7(iv) above; the validity of Peano's fifth postulate, for the particular case of $L1 = N$, was verified in A2.5.9, and the same argument is easily seen to work for general list objects.

(*b*) The reader may have been surprised not to see the third sequent of 4.1.7(iv) (that is, $(\top \vdash_z (\mathsf{iter}_{x,y}(\mathsf{cons}(y, x), [\,], z) = z)))$ appearing as one of the axioms of the τL-calculus. However, now that we have power types available, and hence can formulate Peano's fifth postulate, we may derive this sequent from the other axioms. Specifically, if t denotes the closed term $\{z : LA \mid (\mathsf{iter}_{x,y}(\mathsf{cons}(y, x), [\,], z) = z)\}$, then the first two axioms of 4.3.16(*b*) easily imply the sequents $(\top \vdash ([\,] \in t))$ and $((y \in t) \vdash_{x,y} (\mathsf{cons}(x, y) \in t))$, whence by the Peano postulate we deduce $(\forall y \colon LA)(y \in t)$.

(*c*) In a similar vein, we may show that (the analogues for list objects of) the third and fourth Peano postulates are provable in the τL-calculus; that is, we have $((\mathsf{cons}(x,y) = [\,]) \vdash_{x,y} \bot)$ and

$$((\mathsf{cons}(x,y) = \mathsf{cons}(x',y')) \vdash_{x,x',y,y'} ((x = x') \wedge (y = y'))) .$$

For the first of these, we note that if we substitute $\mathsf{cons}(x,y)$ and $[\,]$ for the variable z in the term $\mathsf{iter}_{u,v}(\widehat{\top}, \widehat{\bot}, z)$, we obtain $\widehat{\top}$ and $\widehat{\bot}$ respectively (where, as in 4.3.2, $\widehat{\phi}$ denotes the term $\{x\colon 1 \mid \phi\}$ of type Ω). So from $(\mathsf{cons}(x,y) = [\,])$ we may deduce $(\widehat{\top} = \widehat{\bot})$, from which we obtain $(* \in \widehat{\bot})$ and hence $\bot$. Similarly, if we substitute $\mathsf{cons}(x,y)$ for z in the term $\mathsf{iter}_{u,v}(v, x'', z)$ we recover x; so from $(\mathsf{cons}(x,y) = \mathsf{cons}(x',y'))$ we may deduce $(x = x')$ (in a context which contains a third variable x'' of type A; but we may then eliminate the latter by the method which we noted at the end of 1.3.1(*a*)). And, if we substitute $\mathsf{cons}(x,y)$ for z in the term $\mathsf{snd}(t)$, where t is the term

$$\mathsf{iter}_{u,v}(\langle \mathsf{cons}(v, \mathsf{fst}(u)), \mathsf{fst}(u)\rangle, \langle [\,], [\,]\rangle, z)$$

of type $LA \times LA$, we recover y. For the derivability of the sequent in (*b*) above enables us to deduce $(\mathsf{fst}(t) = z)$, and hence $(t[\mathsf{cons}(x,y)/z] = \langle \mathsf{cons}(x,y), y\rangle)$.

Given a τL-theory $\mathbb{T}$, we construct its syntactic category $\mathcal{E}_{\mathbb{T}}$ exactly as we did for a τ-theory earlier in this section; the proof that it is a topos of course remains valid. So we need only prove

Proposition 4.3.18 *Let $\mathbb{T}$ be a τL-theory over a signature Σ. Then, for each type A of Σ, the object $\overline{LA}$ of $\mathcal{E}_{\mathbb{T}}$ is a list object over $\overline{A}$, in the sense of* A2.5.15. *In particular, $\mathcal{E}_{\mathbb{T}}$ has list objects.*

Proof The structure maps $o_A\colon \overline{1} \to \overline{LA}$ and $s_A\colon \overline{A} \times \overline{LA} \to \overline{LA}$ are respectively represented by the closed terms $\{z\colon 1 \times LA \mid (\mathsf{snd}(z) = [\,])\}$ and

$$\{z\colon (A \times LA) \times LA \mid (\mathsf{snd}(z) = \mathsf{cons}(\mathsf{fst}(\mathsf{fst}(z)), \mathsf{snd}(\mathsf{fst}(z))))\} .$$

Given an arbitrary object $b\colon PB$ and morphisms $[c]\colon \overline{1} \to b$, $[t]\colon \overline{A} \times b \to b$, we first consider the closed term

$$\begin{aligned} g = \{w\colon P(LA \times B) \mid ((\forall u)((u \in w) \Rightarrow (\mathsf{snd}(u) \in b)) \wedge \\ (\exists y\colon B)((\langle *, y\rangle \in c) \wedge (\langle [\,], y\rangle \in w)) \wedge \\ (\forall z\colon LA)(\forall y\colon B)(\forall x\colon A)((\langle z, y\rangle \in w) \Rightarrow \\ (\exists y'\colon B)((\langle\langle x, y\rangle, y'\rangle \in t) \wedge (\langle \mathsf{cons}(x,z), y'\rangle \in w))))\} \end{aligned}$$

(informally, g is the set of all relations $[w]\colon \overline{LA} \looparrowright b$ which are 'compatible with the recursion data $[c]$, $[t]$'). Then we form the term of type $P(LA \times B)$ which is the intersection of all the members of g, that is

$$f = \{u\colon LA \times B \mid (\forall w\colon P(LA \times B))((w \in g) \Rightarrow (u \in w))\} .$$

Since the relation $\{u\colon LA \times B \mid (\mathsf{snd}(u) \in b)\}$ clearly belongs to g, it is clear that f is (or rather represents) a relation $\overline{LA} \looparrowright b$ in $\mathcal{E}_{\mathbb{T}}$, and that it is compatible with

the recursion data. We claim that it is functional, in the sense defined earlier: for this we just have to show that

$$\{z\colon LA \mid (\exists! y\colon B)(\langle z, y\rangle \in f)\}$$

satisfies the hypotheses of the fifth Peano postulate (where the unique existential quantifier $\exists! y$ is as usual defined by

$$(\exists! y)\phi \equiv (\exists y)(\phi \wedge (\forall y')(\phi[y'/y] \Rightarrow (y = y'))) \;.)$$

But this is a straightforward induction, using the definition of compatibility with the recursion data and the third and fourth Peano postulates as established in 4.3.17(c). A similar induction tells us that $[f]$ is the unique morphism $\overline{LA} \to b$ satisfying the given recursion data; alternatively, we could argue that if $[f']$ is another such, then $(f' \in g)$ must hold, whence we deduce $[f] \leq [f']$ in the allegory $\mathbf{Rel}(\mathcal{E}_{\mathbb{T}})$, and conclude $[f] = [f']$ by A3.2.3.

Thus we have verified that every object in $\mathcal{E}_{\mathbb{T}}$ of the form $\overline{A}$ has a list object. A similar, but more complicated, argument could be given to construct list objects for arbitrary objects of the syntactic category; but we do not need to do so, since we may conclude in particular that $\overline{N} = \overline{L1}$ is a natural number object in $\mathcal{E}_{\mathbb{T}}$, and so we may apply A2.5.17 to obtain arbitrary list objects. □

Remarks 4.3.19 (a) Proposition 4.3.18 implicitly contains a proof of the fact, which we mentioned in Section A2.5, that the five Peano postulates suffice to characterize the natural number object in a topos; for we did not make use of the first two sequents of 4.3.16(b) in the proof, but only of the third and fourth Peano postulates which we deduced from them. However, we shall provide a more explicit proof of this result in 5.1.2 below.

(b) The proof of 4.3.10 now extends in the obvious way to show that, for any τL-theory $\mathbb{T}$, the canonical Σ-structure $M_{\mathbb{T}}$ in $\mathcal{E}_{\mathbb{T}}$ is a conservative $\mathbb{T}$-model. So we obtain the Completeness Theorem for the τL-calculus (i.e. the converse of the Soundness Theorem stated in 4.3.17(a)).

(c) Since, as we observed earlier, any logical functor between toposes preserves list objects, it is clear that a logical functor $F\colon \mathcal{E} \to \mathcal{F}$ (between toposes with natural number objects) induces a functor $\mathbb{T}\text{-Mod}(\mathcal{E}) \to \mathbb{T}\text{-Mod}(\mathcal{F})$, for any τL-theory $\mathbb{T}$. So we have an immediate extension of 4.3.13, proved in exactly the same way: *for any τL-theory $\mathbb{T}$ and any topos $\mathcal{F}$ with a natural number object, the functor $\mathfrak{Log}(\mathcal{E}_{\mathbb{T}}, \mathcal{F}) \to \mathbb{T}\text{-Mod}(\mathcal{F})$ is one half of an equivalence of categories.*

(d) In particular we may, as in 4.3.14(a), consider the empty τL-theory over the empty signature, and thus obtain the *free topos with natural number object*, i.e. the initial object in the full sub-2-category $\mathfrak{Log}_N \subseteq \mathfrak{Log}$ of toposes with natural number objects. (Note that, by A2.5.6(ii), this subcategory is a cosieve in $\mathfrak{Log}$.) And by slicing this topos over a suitable subterminal object as in 4.3.14(e), we may obtain the initial object of the 2-category $\mathfrak{BooLog}_N$ of Boolean toposes with natural number objects. However, the latter is appreciably more complicated than the initial object of $\mathfrak{BooLog}$: in particular, its

unique (up to isomorphism) logical functor to **Set** is not conservative. For in the τL-calculus (indeed, in the τN-calculus) plus the Law of Excluded Middle, we may carry through the proof of Gödel's Incompleteness Theorem; and if ϕ is the unprovable sentence constructed in that proof, then $[\![\phi]\!] \rightarrowtail 1$ is not an isomorphism in the free Boolean topos with natural number object (since ϕ is unprovable), but it is mapped to an isomorphism in **Set**. (And if we slice the free Boolean topos with natural number object over the complement of $[\![\phi]\!]$, we obtain a non-degenerate Boolean topos which admits no logical functor to **Set**.) On the other hand, we shall see in Section F3.3 that the initial object of $\mathfrak{Log}$ is, in a sense, no less complicated than the initial object of $\mathfrak{Log}_N$.

(*e*) Just as in 4.3.15, for any small topos $\mathcal{E}$ with natural number object we may construct a τL-theory $\mathbb{T}_{\mathcal{E}}$ such that $\mathbb{T}_{\mathcal{E}}$-models in $\mathcal{F}$ correspond to logical functors $\mathcal{E} \to \mathcal{F}$. Indeed, we do not need to add any further axioms to the list given in the proof of 4.3.15 – though we could if we wished add function symbols $\hat{\iota}\colon \ulcorner LA \urcorner \to L\ulcorner A \urcorner$ for each type A, together with axioms which say that the interpretation of $\hat{\iota}$ in a model M is a two-sided inverse for the canonical comparison map $F(LA) \to L(FA)$ which exists for any product-preserving functor between categories with list objects. We leave the formulation of these axioms to the reader.

Suggestions for further reading: Bell [95], Boileau & Joyal [138], Fourman [358], Lambek & Scott [682], Lavendhomme & Lucas [688].

D4.4 Predicative type theories

In this section (which may easily be skipped by readers who are anxious to press on with the study of applications of higher-order logic in toposes), we shall attempt to describe the relationship between the ideas of the three preceding sections and the predicative type theories (often associated with the name of P. Martin-Löf, cf. [800]) which are studied by many logicians as a framework for constructive mathematics.

We have seen from an early stage in this book that topos theory is inherently non-classical, in that relatively few toposes satisfy the Law of Excluded Middle (cf. 1.3.3); and we have (particularly in Parts B and C) freely used the word 'constructive' to describe the style of argument which it is appropriate to use when working over a base topos $\mathcal{S}$ which may not be Boolean. However, for many of those who adopt a constructivist view of mathematics, the notion of topos (or the equivalent notion of τ-theory which we described in the last section) is unacceptable, because of the presence of power objects (or power types, as the case may be), and the freedom which they give one to use impredicative definitions.

To explain the meaning of the word 'impredicative' in this context, it is helpful to consider a specific example: one which will be fresh in the mind of anyone who has just read Section D4.3 occurs in the proof of Proposition 4.3.18. Here we

wished to construct a particular morphism $[f]: \overline{LA} \to b$ in our syntactic category $\mathcal{E}_{\mathbb{T}}$; and we did so by considering the intersection of all relations $\overline{LA} \nrightarrow b$ with a particular 'compatibility property'. But $[f]$ itself, once we have constructed it, turns out to be such a relation; in other words, to 'construct' $[f]$, we have quantified over a set which we could not have constructed (in the strict sense of specifying all its members) without constructing $[f]$ itself. (A similar example occurs in our proof of the Knaster–Tarski theorem 4.1.10, where we obtained a fixed point of $f: A \to A$ as the join of a set of 'pre-fixed points' which included the fixed point we were trying to construct.) Although this type of argument is commonplace in the everyday practice of classical mathematicians (it occurs almost everywhere that closure operations are to be found), it is anathema to a strict constructivist. Such people therefore reject the existence of the 'set' which we denoted by the term g in the proof of 4.3.18; and since they do not wish to jettison the Separation Principle that sets may be 'separated out' from pre-existing sets by forming the subset of elements satisfying a given formula, they are bound to reject the existence of the power type $P(LA \times B)$ from which g was separated out by this principle. (For another reason why impredicative power types may be thought 'too powerful' by constructivists, see 4.5.8 below.)

The use of function types (or of list types) does not lead to the same problems of impredicativity; so, if we are searching for a predicative type theory to serve as a foundation for mathematics, a natural starting-point might be the (typed) λN-calculus of Section D4.2. However, this is clearly too weak, since it does not contain any of the first-order logical connectives and quantifiers. There was a good reason for this omission: namely that, as we saw in Section A1.5, a cartesian closed category (even if properly cartesian closed) need not have many of the other 'set-like' properties studied in Chapter A1 – in particular, it need not be regular, and so need not support a sound interpretation even of the regular fragment of first-order logic. But we also saw the remedy for this failing, in A1.5.13: a category which is locally cartesian closed is a Heyting category (and thus supports a sound interpretation of full intuitionistic first-order logic), provided it is also cocartesian. So we might reasonably seek a type calculus which corresponds to locally cartesian closed categories, in the same way that λ-calculus and τ-calculus correspond to cartesian closed categories and toposes respectively. In honour of P. Martin-Löf, who was the first to propose such a type calculus in [799], this calculus is often referred to as *Martin-Löf type theory*; for brevity, we shall call it the *$\mu\lambda$-calculus.*

In this book, we do not have space even to begin to describe the large body of constructive mathematics which has been developed within the $\mu\lambda$-calculus. But, since we have had a good deal to say about locally cartesian closed categories in the book, it seems appropriate to give at least a sketch of how this notion may be translated into type-theoretic terms; and that is our goal in the present section. There are many different ways of presenting the $\mu\lambda$-calculus; the one we have chosen is designed to emphasize its connection with locally cartesian closed categories. However, one feature which they all share, and which differentiates

them from the other calculi we have studied in this chapter, is the fact that the collection of types is no longer 'fixed once for all': in addition to 'constant types', we also have 'variable types' indexed by variables which run over other types, and we shall be able to form 'dependent sums and products' of such types.

Example 4.4.1 To explain these notions, the following simple example (borrowed from [490]) may be helpful. Suppose we wish to define a type whose individuals are dates (of the Christian era, as usually recorded): our first idea might be that this is simply a product type $D \times M \times Y$ (day, month, year), where the individuals of type D are the integers from 1 to 31, those of type M are the integers from 1 to 12, and those of type Y are arbitrary positive integers. But this product type would contain individuals such as 31.9.2001 and 29.2.2002 (in the usual notation), which do not represent actual dates; in order to exclude them, we need to think of D as a variable type $D(m, y)$, depending on variables m and y of types M and Y, and the type of dates is then the dependent sum $\Sigma_{y:Y}\Sigma_{m:M}D(m, y)$. Similarly, the day on which a monthly payment falls due might naively be thought of as an individual of the function type $[M \times Y \to D]$, but it really belongs to the dependent product $\Pi_{y:Y}\Pi_{m:M}D(m, y)$. (Indeed, the type of dates might itself be dependent on variables of other types, such as the country under consideration: for example, 10.9.1752 is a valid date in France, but not in England.)

Our language thus has a new feature, namely *type declarations* which are 'sequents' of the form

$$(\Gamma \vdash (A \in \Sigma\text{-Typ}))$$

where Γ is a context and A is some expression involving free variables $x_1, \ldots, x_n$ which appear in Γ. The type constructors which we used in earlier sections of this chapter become 'rules of inference' for manipulating such type declarations; for example, the presence of binary product types, as in 4.1.1(ii), would correspond to the rule

$$\frac{(\Gamma \vdash (A \in \Sigma\text{-Typ})) \quad (\Gamma \vdash (B \in \Sigma\text{-Typ}))}{(\Gamma \vdash (A \times B \in \Sigma\text{-Typ}))}\ .$$

Moreover, the notion of context becomes more complicated than it has been hitherto (which explains our change of notation from $\vec{x}$ to Γ): a context is now a finite sequence $(x_1 : A_1), (x_2 : A_2), \ldots, (x_n : A_n)$ of distinct *typed* variables, with the proviso that for each $i \leq n$ the type declaration

$$((x_1 : A_1), \ldots, (x_{i-1} : A_{i-1}) \vdash (A_i \in \Sigma\text{-Typ}))$$

is derivable in our system. (So in particular $(A_1 \in \Sigma\text{-Typ})$ must be derivable in the empty context [ ].)

If Γ is a context such that $(\Gamma \vdash (A \in \Sigma\text{-Typ}))$ is derivable, we shall often express this by saying that A is a *type in the context* Γ, or that Γ is a *suitable context* for A. In addition to type declarations, our language will contain *term*

declarations of the form $(\Gamma \vdash (t\colon A))$, where A is a type in the context Γ, and we shall have rules of inference for these too. (The primitive term declarations take the place of function symbols in our language: a function symbol $f\colon A \to B$ becomes a term declaration $((x : A) \vdash (f(x) : B))$. For simplicity, we shall not allow any primitive relation symbols other than equality, since we know from 1.4.9 that we can eliminate them in favour of function symbols.)

Definition 4.4.2 A *$\mu\lambda$-signature* is specified by sets of *basic type declarations* $(\Gamma \vdash (A \in \Sigma\text{-Typ}))$, and *primitive term declarations* $(\Gamma \vdash (t\colon A))$ from which we may derive further declarations by the following rules:

(*a*) the *Weakening Rule*

$$\frac{(\Gamma \vdash \theta)}{(\Delta \vdash \theta)}$$

whenever Γ is a sub-context of Δ (that is, Δ contains all the variables in Γ – possibly not in the same order, but at least in an order that makes it a legitimate context), and θ denotes either an expression of the form $(A \in \Sigma\text{-Typ})$ or one of the form $(t\colon A)$;

(*b*) the *Variable Declaration Rule*

$$\frac{(\Gamma \vdash (A \in \Sigma\text{-Typ}))}{(\Gamma, (x_{n+1} : A) \vdash (x_{n+1} : A))}$$

where x_{n+1} is a variable not appearing in Γ;

(*c*) the *Substitution Rule*

$$\frac{(\Gamma, (y\colon B) \vdash \theta) \quad (\Gamma \vdash (t\colon B))}{(\Gamma \vdash \theta[t/y]))}$$

where, once again, θ may be either a type declaration or a term declaration – but in the latter case the substitution of t for y must be made in both the term and the type;

(*d*) the *Dependent Sum Rules*

$$\frac{(\Gamma, (x\colon A) \vdash (B \in \Sigma\text{-Typ}))}{(\Gamma \vdash (\Sigma_{x:A} B \in \Sigma\text{-Typ}))}$$

(here all occurrences of x in the expression $\Sigma_{x\cdot A} B$ – which we often abbreviate to $\Sigma_x B$, if the type of x is clear from the context – are taken to be bound), and

$$\frac{(\Gamma \vdash (s\colon A)) \quad (\Gamma, (x\colon A) \vdash (B \in \Sigma\text{-Typ})) \quad (\Gamma \vdash (t\colon B[s/x]))}{(\Gamma \vdash (\langle s, t\rangle : \Sigma_{x:A} B))}$$

plus the rules that from $(\Gamma \vdash (t\colon \Sigma_{x\ A} B))$ we may infer $(\Gamma \vdash (\mathsf{fst}(t)\colon A))$ and $(\Gamma \vdash (\mathsf{snd}(t)\colon B[\mathsf{fst}(t)/x]))$;

(*e*) the *Dependent Product Rules*

$$\frac{(\Gamma, (x\colon A) \vdash (B \in \Sigma\text{-Typ})),}{(\Gamma \vdash (\Pi_{x:A} B \in \Sigma\text{-Typ}))} \quad \frac{(\Gamma, (x\colon A) \vdash (t\colon B))}{(\Gamma \vdash (\lambda x\colon A.t\colon \Pi_{x:A} B))}$$

and

$$\frac{(\Gamma \vdash (s\colon \Pi_{x:A} B)) \quad (\Gamma \vdash (t\colon A))}{(\Gamma \vdash (\mathsf{app}(s,t)\colon B[t/x]))}$$

where, once again, the occurrences of x in the expression $\Pi_{x.A} B$ (and in $\lambda x\colon A.t$) are taken to be bound; and

(*f*) the *Unit Type Rules* which are simply the axioms $([\,] \vdash (1 \in \Sigma\text{-Typ}))$ and $([\,] \vdash (*\colon 1))$.

The above rules define the basic notion of a signature for the $\mu\lambda$-calculus. However, for practical purposes, we shall wish to add the formation of equalizers to the type structure in the form of *Equality Types*, for which we have the following rules:

(*g*) Given two term declarations $(\Gamma, (x\colon A) \vdash (s\colon B))$ and $(\Gamma, (x\colon A) \vdash (t\colon B))$, we may infer the type declaration $(\Gamma \vdash (\mathsf{Eq}_x(s,t) \in \Sigma\text{-Typ}))$ (where all occurrences of the variable x in $\mathsf{Eq}_x(s,t)$ are bound) and the term declaration $(\Gamma, (y\colon \mathsf{Eq}_x(s,t)) \vdash (\mathsf{e}(y)\colon A))$.

If we are given a $\mu\lambda$-signature with these extra rules, we shall call it a *$\mu\lambda E$-signature.*

The rules for dependent sums and dependent products are strongly reminiscent of the term constructors for product types and function types, which we introduced in 4.1.2. This is of course no accident, since product and function types are a special case of dependent sums and products: if the type B does not in fact depend on the variable $x\colon A$, then $\Sigma_x B$ will play the rôle of the product type $A \times B$, and $\Pi_x B$ will play the rôle of $[A \to B]$. Thus we do not need to include ordinary (binary) product and function types explicitly in Definition 4.4.2; but we do still need to include the unit type 1.

Note, incidentally, that 4.4.2(*c*) allows us to substitute a term only for the *last* variable in our context (though the limited freedom to permute contexts, allowed to us by (*a*), to some extent compensates for this). The reason is the obvious one that if we substituted for an earlier variable, on which some of the later types in the context were dependent, the removal of that variable from the context would leave something which could not be made into a valid context. For the same reason, dependent sums and products may be indexed only by the last variable in a context.

Before going on to consider the formulae of the $\mu\lambda$-calculus, we next explain how the types and terms are to be interpreted in a locally cartesian closed category $\mathcal{C}$. The basic idea is that a context $\Gamma = (x_1 : A_1), \ldots, (x_n : A_n)$ will be interpreted as a composable string $MA_n \to MA_{n-1} \to \cdots \to MA_2 \to MA_1$ of

morphisms of $\mathcal{C}$. (We shall write $M\Gamma$ for the domain MA_n of the first morphism in the sequence; if Γ is the empty context, we interpret $M\Gamma$ as the terminal object 1 of $\mathcal{C}$.) If B is a type whose existence is derivable in Γ, then it will be interpreted as a further morphism $MB \to M\Gamma$; and any term declaration $(t\colon B)$ in the context Γ will be interpreted as a section of $MB \to M\Gamma$. (Properly, we ought to use some such notation as $M(\Gamma . B)$ for the interpretation of a type B in a context Γ, since the interpretation will definitely depend on Γ as well as B; but this notation is really too cumbersome for practical use.) Formally, we do this as follows:

Definition 4.4.3 Let Σ be a $\mu\lambda$-signature, and $\mathcal{C}$ a locally cartesian closed category. A *structure* M for Σ in $\mathcal{C}$ assigns to each context Γ a composable string of morphisms as above, to each type declaration $(\Gamma \vdash (B \in \Sigma\text{-Typ}))$ a morphism $MB \to M\Gamma$, and to each term declaration $(\Gamma \vdash (t\colon B))$ a morphism $M\Gamma \to MB$ such that the composite $M\Gamma \to MB \to M\Gamma$ is the identity, as follows: the interpretations of the basic type declarations and primitive term declarations are assigned arbitrarily. The interpretations of contexts are defined inductively: if $\Delta = \Gamma, (x_{n+1}\colon A_{n+1})$ and we have already interpreted the context Γ and the type declaration $(\Gamma \vdash (A_{n+1} \in \Sigma\text{-Typ}))$, then the interpretation of Δ is the string $MA_{n+1} \to MA_n \to \cdots \to MA_1$ obtained by concatenating these two interpretations. Derived type and term declarations are interpreted according to the following rules:

(a) For the weakening rule, we deal with two sub-cases: first, when Γ is an initial segment of Δ, and secondly when Δ is obtained from Γ by permuting its terms. In the first case, say $\Gamma = (x_1 : A_1), \ldots, (x_n : A_n)$ and $\Delta = (x_1 : A_1), \ldots, (x_m : A_m)$ where $m > n$, then the composite of the morphisms which lie in the interpretation of Δ but not in that of Γ is a morphism $M\Delta \to M\Gamma$ in $\mathcal{C}$, and we obtain the interpretation of a type or term declaration θ in the context Δ simply by pulling back its interpretation in the context Γ along this morphism. To deal with the case when Δ is obtained by permuting Γ, it suffices to consider that when Δ is obtained by transposing two adjacent members of Γ, say $(x_i : A_i)$ and $(x_{i+1} : A_{i+1})$, since such transpositions generate the full permutation group. Now, since Γ and Δ are both valid contexts, the types A_i and A_{i+1} must both exist in the context $\Theta = (x_1\colon A_1), \ldots, (x_{i-1}\colon A_{i-1})$; so their type declarations in this context are interpreted by morphisms $MA_i \to MA_{i-1}$ and $MA_{i+1} \to MA_{i-1}$, and the interpretation of the type declaration of A_{i+1} in the context $\Theta, (x_i\colon A_i)$ is the pullback of the latter along the former, i.e. the projection $MA_i \times_{MA_{i-1}} MA_{i+1} \to MA_i$. Similarly, the interpretation of the type declaration of A_i in the context $\Theta, (x_{i+1}\colon A_{i+1})$ is the projection $MA_{i+1} \times_{MA_{i-1}} MA_i \to MA_{i+1}$. But the domains of these two projections are (canonically) isomorphic, and this isomorphism now 'propagates up' the interpretations of the rest of Γ and Δ to yield a canonical isomorphism $M\Delta \cong M\Gamma$. So we obtain the interpretation of θ in the new context Δ from its interpretation in Γ by

composing with this isomorphism (or its inverse, depending on whether θ is a type or a term declaration).

(*b*) For the variable declaration rule, if we have interpreted the type declaration of A_{n+1} in the context Γ as a morphism $MA_{n+1} \to M\Gamma = MA_n$, then its interpretation in the extended context $\Gamma, (x_{n+1} : A_{n+1})$ is the projection $MA_{n+1} \times_{MA_n} MA_{n+1} \to MA_{n+1}$, and we interpret the variable declaration $(x_{n+1} : A_{n+1})$ in this context as the diagonal map.

(*c*) The substitution rule is simply interpreted by pullback: if the term declaration $(t : B)$ has been interpreted in the context Γ as a morphism $M\Gamma \to MB$, then we obtain the interpretation of $\theta[t/y]$ in this context from that of θ in the extended context $\Gamma, (y : B)$ by pulling back along this morphism. (Note that, if the type and/or term appearing in θ does not actually involve the variable y – that is, if it already exists in the context Γ – we shall arrive back at its original interpretation in that context if we add a 'dummy variable' y and then substitute a term for it.)

(*d*–*e*) The dependent sum and product types $\Sigma_{x:A}B$ and $\Pi_{x:A}B$ are of course interpreted (in a context Γ) by applying the left and right adjoints to pullback along $MA \to M\Gamma$ to the interpretation of B in the extended context $\Gamma, (x : A)$. For the terms associated with these types, we proceed as follows:

(i) The interpretation of $(\langle s, t\rangle : \Sigma_x B)$ in the context Γ is the top composite in the diagram

$$\begin{tikzcd} M\Gamma \arrow[r, "Mt"] & M(B[s/x]) \arrow[r] \arrow[d] & MB \arrow[d] \\ & M\Gamma \arrow[r, "Ms"] & MA \end{tikzcd}$$

where Ms and Mt are the interpretations of s and t in the context Γ, and the square is a pullback. (Note that the interpretation of $\Sigma_x B$ in the context Γ is simply the composite $MB \to MA \to M\Gamma$, so this makes sense.)

(ii) Given the interpretation of a term $(t : \Sigma_x B)$ in Γ, the interpretation of $(\mathsf{fst}(t) : A)$ is the composite of this morphism with the morphism $MB \to MA$ which interprets the type B in the context $\Gamma, (x : A)$, and the interpretation of $(\mathsf{snd}(t) : B[\mathsf{fst}(t)/x])$ is its factorization through the pullback which is used to define the interpretation of $B[\mathsf{fst}(t)/x]$ in Γ.

(iii) The interpretation of $(\lambda x\,.\,t : \Pi_x B)$ in the context Γ is obtained by regarding the interpretation of $(t : B)$ as a morphism $f^*(1_{M\Gamma}) \cong 1_{MA} \to MB$ in $\mathcal{C}/MA$ (where f is the name of the morphism $MA \to M\Gamma$ interpreting the type A), and transposing it across the adjunction $(f^* \dashv \Pi_f)$.

(iv) Given the interpretations Ms, Mt of terms $(s : \Pi_x B)$ and $(t : A)$ in the context Γ, we interpret $(\mathsf{app}(s, t) : B[t/x])$ as the pullback along Mt of the

composite

$$MA \xrightarrow{f^*(Ms)} f^*M\Pi_x B \longrightarrow MB$$

in $\mathcal{C}/MA$, where f, as before, is the name of the morphism $MA \to M\Gamma$ and the second factor above is the counit of $(f^* \dashv \Pi_f)$.

(f) The unit type 1 is of course interpreted in the empty context as the terminal object of $\mathcal{C}$, and the term $(*\colon 1)$ is interpreted as the identity morphism on this object.

(g) If Σ has equality types as in 4.4.2(g), then $M\mathsf{Eq}_x(s,t)$ is interpreted in a context Γ as the composite $E \rightarrowtail MA \to M\Gamma$, where $E \rightarrowtail MA$ is the equalizer of the morphisms $MA \rightrightarrows MB$ interpreting s and t, and the term $\mathsf{e}(y) : A$ is interpreted in the context $\Gamma, (y\colon \mathsf{Eq}_x(s,t))$ as the morphism $E \to E \times_{M\Gamma} MA$ which is the graph of $E \rightarrowtail MA$ regarded as a morphism in $\mathcal{C}/M\Gamma$.

Now we may turn to considering the formulae of the $\mu\lambda$-calculus. Of course, our atomic formulae are simply equalities-in-context $\Gamma.(s =_A t)$, where s and t are terms such that the term declarations $(\Gamma \vdash (s : A))$ and $(\Gamma \vdash (t : A))$ are derivable in Σ. We shall distinguish between the *full* $\mu\lambda$-calculus, in which we allow ourselves the full range of (finitary) first-order logical connectives and quantifiers from 1.1.3 – but with the restriction that we may quantify only over the *last* variable in a context (though we are of course free to permute the variables in a context, provided it remains a valid context) – and the *cartesian* $\mu\lambda$-calculus in which we allow only cartesian logic as defined in 1.3.4. If ϕ and ψ are formulae, we now write the sequent asserting that ψ is a logical consequence of ϕ in the context Γ as $(\Gamma, \phi \vdash \psi)$ – though if ϕ is the constant $\top$, as it very often will be, we omit it from this notation.

Definition 4.4.4 Given a structure M for a $\mu\lambda$-signature Σ in a locally cartesian closed category $\mathcal{C}$, we assign to each cartesian formula-in-context $\Gamma.\phi$ an *interpretation* $[\![\Gamma.\phi]\!]_M \rightarrowtail M\Gamma$ in a straightforward way; and if $\mathcal{C}$ is also a Heyting category, then we may extend this interpretation to arbitrary first-order formulae-in-context. Explicitly, an atomic formula $\Gamma.(s =_A t)$ is interpreted as the equalizer of the morphisms $M\Gamma \rightrightarrows MA$ interpreting s and t in the context Γ, the first-order connnectives are interpreted exactly as in 1.2.6, and quantifiers in almost exactly the same way – the only difference being that the morphism of $\mathcal{C}$ along which we apply the left and right adjoints to pullback is no longer necessarily a product projection. We say that a sequent $(\Gamma, \phi \vdash \psi)$ is *satisfied* in a structure M if $[\![\Gamma.\phi]\!]_M \leq [\![\Gamma.\psi]\!]_M$ in $\mathrm{Sub}(M\Gamma)$. And if $\mathbb{T}$ is a $\mu\lambda$-theory (that is, a set of sequents) over Σ, we say M is a *model* of $\mathbb{T}$ if it satisfies all the sequents in $\mathbb{T}$.

Definition 4.4.5 The *rules of inference* of the full (resp. cartesian) $\mu\lambda$-calculus are all the finitary first-order rules of 1.3.1(a–g) (resp. the rules of regular logic) with substitution restricted to the last variable in a context, but with weakening

allowed in all cases where it makes sense, plus the following rules for dependent sums and products, and the unit type:

(*a*) If $(\Gamma \vdash (t\colon \Sigma_{x:A} B))$, then we may infer $(\Gamma \vdash (t = \langle \mathsf{fst}(t), \mathsf{snd}(t)\rangle))$. Similarly, if $(\Gamma \vdash (s : A))$ and $(\Gamma \vdash (t : B[s/x]))$, then $(\Gamma \vdash (\mathsf{fst}(\langle s,t\rangle) = s))$ and $(\Gamma \vdash (\mathsf{snd}(\langle s,t\rangle) = t))$.

(*b*) If $(\Gamma \vdash (t\colon A))$ and $(\Gamma, (x : A) \vdash (s : B))$, then $(\Gamma \vdash (\mathsf{app}(\lambda x \,.\, s, t) =_{B[t/x]} s[t/x]))$. And if $(\Gamma \vdash (t\colon \Pi_{x:A} B))$, then $(\Gamma \vdash (t = \lambda x\colon A \,.\, \mathsf{app}(t, x)))$.

(*c*) For the unit type, we have the single axiom $((x\colon 1) \vdash (x =_1 *))$.

If we are working over a $\mu\lambda E$-signature, then we adopt the following additional rules for equality types:

(*d*) We have the axiom

$$(\Gamma, (y\colon \mathsf{Eq}_x(s,t)), (y'\colon \mathsf{Eq}_x(s,t)), (\mathsf{e}(y) = \mathsf{e}(y')) \vdash (y = y'))$$

and the rule that, from the term declarations $(\Gamma \vdash (r : A))$, $(\Gamma, (x : A) \vdash (s\colon B))$ and $(\Gamma, (x\colon A) \vdash (t\colon B))$ plus the equation $(\Gamma \vdash (s[r/x] = t[r/x]))$, we may infer a new term declaration $(\Gamma \vdash (\mathsf{fact}_{s,t}(r)\colon \mathsf{Eq}_x(s,t)))$ and the equation $(\Gamma \vdash (r = \mathsf{e}(\mathsf{fact}_{s,t}(r))))$. (We normally omit the subscripts from the term $\mathsf{fact}_{s,t}(r)$.)

It will be noted that the rules of 4.4.5(*a*–*c*) are exactly those of 4.2.3(*c*–*d*), adapted to our new setting. In particular, the proof that the rules of 4.2.3 are valid in any cartesian closed category easily generalizes to show that those of 4.4.5 are valid in any locally cartesian closed category; and since the first-order rules are valid in any Heyting category, we obtain the usual Soundness Theorem:

Proposition 4.4.6 *Let $\mathbb{T}$ be a cartesian (resp. full) $\mu\lambda$-theory over a $\mu\lambda$-signature Σ, and let σ be a sequent over Σ. If σ is derivable in $\mathbb{T}$, then it is satisfied in all models of $\mathbb{T}$ in locally cartesian closed categories (resp. locally cartesian closed Heyting categories).* □

It is clear, too, that 4.4.6 extends straightforwardly to theories over $\mu\lambda E$-signatures: the first axiom of 4.4.5(*d*) is valid because equalizers are monic, and if we interpret the term $\mathsf{fact}_{s,t}(r)$ by the unique factorization of Mr through the equalizer of Ms and Mt, then the rule is easily seen to be sound.

The reason for adding equality types to our signature becomes apparent when we consider the problem of building a syntactic category for a $\mu\lambda$-theory. The problem is that, if we proceed along the lines of Section D1.4 (that is, if we take the objects of the category to be (the denotations of) formulae-in-context $(x\colon A)\,.\,\phi)$, we are faced with the requirement to construct exponentials of the form $(y\colon B)\,.\,\psi^{(x.A)\,\phi}$, and (in the absence of power types) we cannot readily construct this from the particular case $\phi = \psi = \top$, which we have already 'built into' the type structure. So, as in Section D4.2, we must take the objects of our

syntactic category to be types, and therefore we need to have equalizers, as well as finite products, built into the type structure as well.

Explicitly, given a cartesian $\mu\lambda$-theory $\mathbb{T}$ over a $\mu\lambda E$-signature Σ, the objects of our syntactic category $\mathcal{C}_{\mathbb{T}}$ will be the types of Σ (in the empty context). If A and B are such types, then morphisms $A \to B$ will be represented by terms t for which the term declaration $((x\colon A) \vdash (t\colon B))$ is derivable, two such terms t and t' representing the same morphism if $((x : A) \vdash (t = t'))$ is derivable in $\mathbb{T}$. Composition is defined by substitution, as in Section D4.2; the identity morphisms of $\mathcal{C}_{\mathbb{T}}$ are provided by the variable declaration rule of 4.4.2(b).

Proposition 4.4.7 *The category $\mathcal{C}_{\mathbb{T}}$ defined above is locally cartesian closed.*

Proof First we show that $\mathcal{C}_{\mathbb{T}}$ is properly cartesian closed; except for the construction of equalizers, this is exactly like the proof of 4.2.5. The terminal object is the type 1, and the unique morphism $A \to 1$ is represented by $((x\colon A) \vdash (*\colon 1))$, which we obtain from 4.4.2(f) by weakening. The product of two types A and B is $\Sigma_{x.A}B$ (where we use the weakening rule to regard B as a type in the context $(x : A)$, of course); the product projections are induced by fst and snd in the obvious way, and morphisms into the product are induced by the term constructor $\langle -, - \rangle$. The equalizer of two morphisms $[((x : A) \vdash (s : B))]$ and $[((x : A) \vdash (t : B))]$ is of course given by the equality type $\mathsf{Eq}_x(s, t)$, together with the morphism $[((y\colon \mathsf{Eq}_x(s,t)) \vdash (\mathsf{e}(y)\colon A))]$. And the exponential B^A is the dependent product $\Pi_{x:A}B$, with structure induced by app and $\lambda x.-$ as in the proof of 4.2.5.

Next, we observe that, if A is a particular type in the empty context, we may similarly build a category $\mathcal{C}_{\mathbb{T}}[A]$ whose objects are types in the context $(x\colon A)$, and (since all the rules of the calculus are 'insensitive' to adding an additional term to the beginning of a context) the same arguments will show that $\mathcal{C}_{\mathbb{T}}[A]$ is (properly) cartesian closed. We claim that $\mathcal{C}_{\mathbb{T}}[A]$ is equivalent to the slice category $\mathcal{C}_{\mathbb{T}}/A$.

To prove this, observe first that we have a functor $\Phi\colon \mathcal{C}_{\mathbb{T}}[A] \to \mathcal{C}_{\mathbb{T}}/A$ which sends a type B in the context $(x : A)$ to $\Sigma_{x\,A}B$, equipped with the morphism to A represented by the term declaration $((z\colon \Sigma_x B) \vdash (\mathsf{fst}(z)\colon A))$. A morphism $B \to C$ in $\mathcal{C}_{\mathbb{T}}[A]$, represented by a term declaration $((x\colon A), (y\colon B) \vdash (t\colon C))$, is mapped to the morphism represented by

$$((z\colon \Sigma_x B) \vdash (\langle \mathsf{fst}(z), t[\mathsf{fst}(z), \mathsf{snd}(z)/x, y]\rangle \colon \Sigma_x C))\ .$$

It is easy to check that this is indeed functorial. In the opposite direction, the functor $\Psi\colon \mathcal{C}_{\mathbb{T}}/A \to \mathcal{C}_{\mathbb{T}}[A]$ sends an object B of $\mathcal{C}_{\mathbb{T}}$ equipped with a morphism $[((y\colon B) \vdash (t\colon A))]\colon B \to A$ to the equality type $\mathsf{Eq}_y(t, x)$ in the context $(x\colon A)$; and if $[r]\colon [t] \to [s]$ is a morphism of $\mathcal{C}_{\mathbb{T}}/A$, we map it to the morphism represented by the term declaration

$$((w\colon \mathsf{Eq}_y(t,x)) \vdash (\mathsf{fact}_{s,x}(r[\mathsf{e}(w)/y])\colon \mathsf{Eq}_z(s,x)))\ .$$

(Note that, as usual, the construction of the functor in this direction depends on our ability to choose representatives for the morphisms in $\mathcal{C}_{\mathbb{T}}$, as in 1.4.3.) Now, given a type B in the context $(x\colon A)$, we have an isomorphism between B and $\Psi\Phi(B)$ represented by the term declarations

$$((x\colon A),(y\colon B) \vdash (\mathsf{fact}(\langle x, y\rangle)\colon \mathsf{Eq}_z(\mathsf{fst}(z), x)))$$

and

$$((x\colon A),(w\colon \mathsf{Eq}_z(\mathsf{fst}(z), x)) \vdash (\mathsf{snd}(\mathsf{e}(w))\colon B[x/x]))\ .$$

It is straightforward to verify that this is a natural isomorphism between $\Psi\Phi$ and the identity; and the construction of a natural isomorphism $\Phi\Psi \cong 1$ is similar. So the claim is established, and hence $\mathcal{C}_{\mathbb{T}}$ is locally cartesian closed. □

Remark 4.4.8 In the same way that we defined the category $\mathcal{C}_{\mathbb{T}}[A]$ in the proof above, we could of course have defined a category $\mathcal{C}_{\mathbb{T}}[\Gamma]$ for any context Γ, and proved that it is cartesian closed. If Γ is the context $(x_1 : A_1), \ldots, (x_n : A_n)$, this category may be shown to be equivalent to the slice category $\mathcal{C}_{\mathbb{T}}/\Sigma_{x_1\ A_1}\Sigma_{x_2.A_2}\cdots\Sigma_{x_{n-1}:A_{n-1}}A_n$.

It is now straightforward as usual to define a Σ-structure in the category $\mathcal{C}_{\mathbb{T}}$, and to prove that it satisfies exactly those cartesian sequents which are derivable in $\mathbb{T}$. So we obtain the expected completeness theorem:

Theorem 4.4.9 *Let Σ be a $\mu\lambda E$-signature, $\mathbb{T}$ a cartesian $\mu\lambda$-theory over Σ, and σ a cartesian sequent over Σ. Then σ is satisfied in all $\mathbb{T}$-models in locally cartesian closed categories iff it is derivable in $\mathbb{T}$.* □

If we wish to put back the full first-order logic, the simplest way of doing so is to rely on A1.5.13, and to ensure that our syntactic category has enough cocartesian structure (finite coproducts and coequalizers of effective equivalence relations) to imply that it is a Heyting category. We may do this by building yet more type constructors into our signature: in fact we choose to add disjoint coproducts and coequalizers of arbitrary equivalence relations, so that the corresponding class of categories is actually that of locally cartesian closed pretoposes. (The construction of quotients by equivalence relations, though it uses the power-set axiom in the traditional set-theoretic approach, is predicative and hence acceptable to 'strict' constructivists.) We have room here for only a brief sketch of how this may be done; for a much more detailed account, we refer the reader to [490].

For coproducts, we first assume given a particular type 2 (in the empty context) with closed terms $\dagger, \ddagger$ of type 2, for which the sequent $((\dagger = \ddagger) \vdash \bot)$ is an axiom (in the empty context). (Note that we do not have to add the empty coproduct 0 explicitly, since we can obtain it as the equality type $\mathsf{Eq}_x(\dagger, \ddagger)$ where

x is of type 1.) Then we introduce a new type constructor

$$\frac{(\Gamma \vdash (A \in \Sigma\text{-Typ})) \quad (\Gamma \vdash (B \in \Sigma\text{-Typ}))}{(\Gamma, (x: 2) \vdash ((\mathsf{case}\ x\ \mathsf{of}\ A\ \mathsf{or}\ B) \in \Sigma\text{-Typ}))}$$

with the convention that this type reduces to A (resp. B) if † (resp. ‡) is substituted for x, and a similar constructor for terms: given terms s, t of types A, B in a context Γ, we have a term ($\mathsf{case}\ x\ \mathsf{of}\ s\ \mathsf{or}\ t$) of type ($\mathsf{case}\ x\ \mathsf{of}\ A\ \mathsf{or}\ B$) in the context $\Gamma, (x: 2)$. The coproduct of A and B in the syntactic category will then be $\Sigma_{x:2}(\mathsf{case}\ x\ \mathsf{of}\ A\ \mathsf{or}\ B)$. (Of course, we need further term constructors and axioms to ensure that this is indeed a coproduct; but we shall not describe them in detail here.)

For coequalizers, our type constructor starts from the assumption that we have type declarations $(\Gamma \vdash (A \in \Sigma\text{-Typ}))$ and $(\Gamma \vdash (R \in \Sigma\text{-Typ}))$, plus two terms $s, t: A$ in the context $\Gamma, (y: R)$ for which the sequents asserting that $R \rightrightarrows A$ is an equivalence relation (in the category $\mathcal{C}_{\mathbb{T}}[\Gamma]$) are derivable. Then we may infer a new type declaration $(\Gamma \vdash (A/R \in \Sigma\text{-Typ}))$ and a term declaration $(\Gamma, (x: A) \vdash (\mathsf{q}(x) : A/R))$ satisfying appropriate axioms. To ensure that $[\mathsf{q}(x)]: A \to A/R$ really is a coequalizer in $\mathcal{C}_{\mathbb{T}}[\Gamma]$, we need a further term constructor which, given a term r of type C in the context $\Gamma, (x: A)$ for which the sequent

$$(\Gamma, (y: R) \vdash (r[t/x] = r[s/x]))$$

is derivable, produces a new term $\mathsf{cofact}_{s,t}(r, w)$ in which the occurrences of x in R are bound, but which has a new free variable w of type A/R. We leave to the reader the task of formulating the axioms which say that $[\mathsf{q}(x)]$ is a coequalizer and $([s], [t])$ is its kernel-pair.

On adding these type and term constructors to the (full) $\mu\lambda E$-calculus, we arrive at a logical calculus which we might call the *$\mu\lambda F$-calculus* (F for 'full first-order logic'), which is interpretable in arbitrary cartesian closed pretoposes and which satisfies a completeness theorem for interpretations in such categories. Moreover, unlike the τ-calculus which we studied in the previous section, it is acceptable to strict constructivists. However, as will be clear from what we have done in this section, that acceptability has been won at the cost of a very considerable increase in the complexity of the calculus itself, and so we shall not pursue it further.

Suggestions for further reading: Beeson [90], Jacobs [490], Maietti [772], Maietti & Valentini [774], Martin-Löf [799, 800], Sambin [1079, 1080], Seely [1107].

D4.5 Axioms of choice and booleanness

For the remaining three sections of this chapter, we turn from the formal business of interpreting higher-order logic in categories such as toposes to its exploitation: that is, to the consideration of certain properties of toposes that are best

studied by making use of the higher-order internal logic of toposes. (This theme will continue in Chapter D5, where we look again at natural number objects, and at various notions of finiteness in toposes.) We begin by considering the rôle of the axiom of choïce in topos theory.

It is well known that the classical axiom of choice is equivalent to the statement that every surjective map is a split epimorphism in the category **Set**. Accordingly, the assertion 'every epimorphism splits' is often taken as the correct formulation of the axiom of choice for a general topos. However, this statement is really the conjunction of two conditions, one of which holds for **Set** even if we do not assume the axiom of choice. In a general topos, it is frequently useful to consider these two conditions separately.

Definition 4.5.1

(*a*) We say that *supports split* in a topos $\mathcal{E}$, or that $\mathcal{E}$ satisfies (SS), if every subterminal object is projective in $\mathcal{E}$; equivalently, if for every object A the epic part of the image factorization of $A \to 1$ is split epic.

(*b*) We say that an object A of a topos $\mathcal{E}$ is *internally projective* if $\Pi_A : \mathcal{E}/A \to \mathcal{E}$ is a regular functor (equivalently, preserves epimorphisms). And we say $\mathcal{E}$ satisfies the *internal axiom of choice* (IC) if every object of $\mathcal{E}$ is internally projective.

Examples 4.5.2 (*a*) If $\mathcal{E}$ is the topos $[\alpha^{\mathrm{op}}, \mathbf{Set}]$ where α is an ordinal (regarded as a well-ordered set), then it is easy to see that every nonzero subterminal object is representable, and hence projective. So $\mathcal{E}$ satisfes (SS); but if $\alpha > 1$ then $\mathcal{E}$ is not Boolean, from which it follows by 4.5.8 below that it cannot satisfy (IC).

(*b*) If G is a group, then the topos $[G, \mathbf{Set}]$ satisfies (IC) provided **Set** does; for the forgetful functor $[G, \mathbf{Set}] \to \mathbf{Set}$ is logical, and so commutes with the functors Π_A. However, if G has more than one element then $[G, \mathbf{Set}]$ does not satisfy (SS), since the epimorphism $G \twoheadrightarrow 1$ (where G acts on itself by left translations) fails to split.

(*c*) If G is a topological group, then the topos $\mathbf{Cont}(G)$ of continuous G-sets may fail to satisfy (IC), even if we assume it in **Set**. For example, if we take $G = \mathbb{Z}$ with the topology generated by the filter of subgroups of finite index, $A = \mathbb{N}$ with trivial G-action, and B to be the coproduct of the finite cyclic G-sets C_n of all possible sizes, then $\Pi_A(B \to A)$ is the coreflection in $\mathbf{Cont}(G)$ of the product $\Pi_{n \in \mathbb{N}} C_n$ as computed in $[G, \mathbf{Set}]$; but all G-orbits of the latter are infinite, so its coreflection is empty and does not map epimorphically to $1 = \Pi_A(1_A)$.

(*d*) Let X be a locale. An open sublocale U of X, regarded as a subterminal object of $\mathbf{Sh}(X)$, is projective iff every open covering of U has a pairwise-disjoint refinement. For, if this condition holds and we are given a local homeomorphism $E \to X$ such that $E \to U$ is epic, then we can cover U by opens over which E admits a section, and after refining this cover to a pairwise-disjoint one we can patch the sections together. Conversely, if U is projective and $(V_i \mid i \in I)$ is an open cover of U, then the disjoint union E of the V_i is the domain of a

local homeomorphism over X with support U, and a section $s: U \to E$ defines a pairwise-disjoint cover of U by the sublocales $s^{-1}(V_i) \subseteq V_i$. Thus the topos $\mathbf{Sh}(X)$ satisfies (SS) iff every family of open sublocales of X can be refined to a pairwise-disjoint family with the same union: such locales are sometimes called *strongly zero-dimensional.* We note that they include all zero-dimensional locales X which are 'second countable', i.e. such that $\mathcal{O}(X)$ is countably generated as a complete join-semilattice. For, given a family $\mathcal{U}$ of opens in such a locale, we may refine it to a countable family $\mathcal{V} = \{V_1, V_2, V_3, \ldots\}$ of clopens with the same union; then, if we define $W_1 = V_1$, $W_2 = V_2 \cap \complement V_1$, $W_3 = V_3 \cap \complement(V_1 \cup V_2)$, $\ldots$, we obtain a pairwise-disjoint family with the same union as $\mathcal{V}$ (and hence as $\mathcal{U}$). (On the other hand, $\mathbf{Sh}(X)$ satisfies (IC) iff it satisfies (AC), iff $\mathcal{O}(X)$ is Boolean; cf. 4.5.15 below.)

(*e*) We remark that, even for spatial locales X, the condition of second countability in the previous paragraph cannot be weakened to separability (the possession of a countable dense subset). For a counterexample, consider the 'Sorgenfrey plane', that is the set $\mathbb{R}^2$ with the topology generated by all 'half-open rectangles' $[a, b) \times [c, d)$ $(a, b, c, d \in \mathbb{R})$. This is clearly zero-dimensional, since the complements of all such rectangles are also open; and it is separable since $\mathbb{Q} \times \mathbb{Q}$ is dense in it. But the family of open sets

$$\{[a, a+1) \times [-a, -a+1) \mid a \in \mathbb{R}\}$$

has no pairwise-disjoint refinement. For if $\mathcal{U}$ were a pairwise-disjoint family of open sets, each contained in a member of the above family, then for each $\epsilon > 0$ the set

$$\{a \in \mathbb{R} \mid [a, a+\epsilon) \times [-a, -a+\epsilon) \text{ is contained in a member of } \mathcal{U}\}$$

would be countable, and hence only countably many points of the line $\{\langle a, -a\rangle \mid a \in \mathbb{R}\}$ could belong to members of $\mathcal{U}$.

In order to establish the relationship between the conditions of 4.5.1 and the axiom of choice, we need

Lemma 4.5.3 *For an object A of a topos $\mathcal{E}$, the following are equivalent:*

(i) *A is internally projective.*
(ii) *The functor $(-)^A: \mathcal{E} \to \mathcal{E}$ preserves epimorphisms.*
(iii) *For every epimorphism $e: B \twoheadrightarrow A$, there exists $C \twoheadrightarrow 1$ such that $C^*(e)$ is split epic.*

Proof (i) $\Rightarrow$ (ii) since $(-)^A$ is the composite $\Sigma_A A^*$, and Σ_A preserves epimorphisms.

(ii) $\Rightarrow$ (i) because we have a pullback square

$$\begin{tikzcd} \Pi_A(f) \arrow[r] \arrow[d] & B^A \arrow[d, "f^A"] \\ 1 \arrow[r, "\overline{1_A}"] & A^A \end{tikzcd}$$

for any $f: B \to A$, and pullback along a fixed morphism preserves epimorphisms.

(i) $\Rightarrow$ (iii): Given e, take C to be $\Pi_A(e)$, which maps epimorphically to $\Pi_A(1_A) \cong 1$. The counit map $A^*(C) \to e$ in $\mathcal{E}/A$ can be regarded as a splitting of $C^*(e)$ in $\mathcal{E}/C$.

(iii) $\Rightarrow$ (i) since C^* commutes with the functors Π_A in an appropriate sense, and reflects epimorphisms. □

Corollary 4.5.4 *For a topos $\mathcal{E}$, the following are equivalent:*

(i) *Every object of $\mathcal{E}$ is projective.*

(ii) *Every epimorphism in $\mathcal{E}$ is split.*

(iii) *$\mathcal{E}$ satisfies* (SS) *and* (IC).

(iv) *$\mathcal{E}$ satisfies* (IC) *and its terminal object is projective.*

Proof (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) is trivial.

For (iv) $\Rightarrow$ (ii), (IC) says that every epimorphism e in $\mathcal{E}$ acquires a splitting when pulled back along some $C \twoheadrightarrow 1$, by 4.5.3; but if $C \twoheadrightarrow 1$ itself has a splitting $c: 1 \to C$, we can pull back the splitting of C^*e along it to obtain a splitting for e. □

A topos satisfying the equivalent conditions of 4.5.4 will be said to satisfy the (*external*) *axiom of choice* (AC).

The usual set-theoretic formulation of the axiom of choice is concerned not with projective objects but with *choice objects*, that is objects A admitting a choice function $c: P^+A \to A$, where P^+A is the object of inhabited subobjects of A (that is, the image of $\in_A \rightarrowtail PA \times A \to PA$), such that the sentence $(\forall w: P^+A)(c(w) \in w)$ is satisfied – equivalently, the graph of c factors through $\in_A \rightarrowtail P^+A \times A$. To establish the relationship between choiceness and projectivity, we need to consider entire relations: we say a relation $\phi: A \nrightarrow B$ (in the sense of Section A3.1) in a regular category is *entire* if $\phi^\circ\phi \geq 1_A$.

Lemma 4.5.5 *An object A of a regular category is projective (with respect to covers) iff every entire relation with domain A contains (the graph of) a morphism.*

Proof Let $\phi: A \nrightarrow B$ be a relation. If $(u: T \to A, v: T \to B)$ is a tabulation of ϕ, then the assertion that ϕ is entire says precisely that u is a cover. If s is a

splitting of it, then the morphism $f = vs$ clearly satisfies $f_\bullet \leq \phi$. Conversely, if A satisfies the condition, let $e: B \twoheadrightarrow A$ be a cover. Then $e^\bullet: A ⇸ B$ is an entire relation; so we have $f: A \to B$ with $f_\bullet \leq e^\bullet$ and hence $(ef)_\bullet \leq e_\bullet e^\bullet = 1_A$, whence $ef = 1_A$ by A3.2.3(ii). □

Lemma 4.5.6 *An object A of a topos is choice iff every entire relation with codomain A contains a morphism.*

Proof To say that $\phi: B ⇸ A$ is entire means precisely that $\ulcorner\phi\urcorner: B \to PA$ factors through $P^+A \rightarrowtail PA$. If we are given a choice function c, then the morphism $f = c\ulcorner\phi\urcorner$ has the required properties. For the converse, we consider $\in_A$ itself as a relation $P^+A ⇸ A$; to say that a morphism $P^+A \to A$ is contained in this relation is exactly to say that it is a choice function. □

Thus we see that, in a topos, (AC) is equivalent to the assertion that every object is choice. There is also a notion of 'internal choiceness': A is internally choice if the object of choice functions for A, constructed as an appropriate subobject of A^{P^+A}, is well-supported. And the above argument, with pullback functors C^* inserted at the appropriate places, shows that $\mathcal{E}$ satisfies (IC) iff every object of $\mathcal{E}$ is internally choice. Note, however, that A is internally choice iff there exists $C \twoheadrightarrow 1$ such that C^*A is a choice object in $\mathcal{E}/C$; in fact we can take C to be the object of choice functions for A. Thus there is not much gain in generality by working with internally rather than externally choice objects, and from now on we shall mainly confine ourselves to the latter.

In any category, a finite coproduct of projective objects is projective. For choice objects, the picture is distinctly different:

Proposition 4.5.7 *The following conditions on a topos $\mathcal{E}$ are equivalent:*

(i) *$\mathcal{E}$ is Boolean.*

(ii) *Finite coproducts of choice objects are choice.*

(iii) *$2 = 1 \amalg 1$ is a choice object.*

(iv) *2 is an internally choice object.*

Proof (i) ⇒ (ii): Suppose A_1 and A_2 are choice objects, and let

$$B \xleftarrow{\;u\;} T \xrightarrow{\;v\;} A_1 \amalg A_2$$

be a tabulation of an entire relation $\phi: B ⇸ A_1 \amalg A_2$. Let $T_1 = v^*(A_1)$, let B_1 be the image of $T_1 \rightarrowtail T \to B$, let B_2 be the complement of B_1 in $\mathrm{Sub}(B)$, and let $T_2 = u^*(B_2)$. Then $T_2 \cap T_1 = 0$ in $\mathrm{Sub}(T)$, so v maps T_2 into A_2; and T_1 and T_2 tabulate entire relations $\phi_1: B_1 ⇸ A_1$ and $\phi_2: B_2 ⇸ A_2$ respectively. By assumption, these contain morphisms $f_1: B_1 \to A_1$ and $f_2: B_2 \to A_2$; but $B \cong B_1 \amalg B_2$, so these combine to yield a morphism $f: B \to A_1 \amalg A_2$, which is contained in the relation tabulated by $T_1 \amalg T_2$ and hence in ϕ.

(ii) $\Rightarrow$ (iii): 1 is trivially choice, since any entire relation with codomain 1 is already a morphism (cf. A3.2.8); so (ii) implies that $2 = 1 \amalg 1$ is choice.

(iii) $\Rightarrow$ (iv) is trivial.

(iv) $\Rightarrow$ (i): Clearly, if there exists $C \twoheadrightarrow 1$ such that $\mathcal{E}/C$ is Boolean, then $\mathcal{E}$ is Boolean, since C^* reflects the validity of the identity $\neg\neg = 1_\Omega$. So we may as well assume that 2 is (externally) choice. It further suffices to prove that $\mathrm{Sub}_{\mathcal{E}}(1)$ is Boolean, since the hypothesis that 2 is choice is preserved by $B^* \colon \mathcal{E} \to \mathcal{E}/B$ for any B. So let U be a subterminal object in $\mathcal{E}$, and form the cokernel-pair $1 \rightrightarrows B$ of $U \rightarrowtail 1$. As in the proof of 4.5.5, we may regard the epimorphism $1 \amalg 1 \twoheadrightarrow B$ as an entire relation $B \looparrowright 2$, and hence find a splitting for it; so B may be identified with a subobject of 2, and is therefore decidable by A1.4.15. But $\mathcal{E}^{\mathrm{op}}$ is regular, so $U \rightarrowtail 1$ is the equalizer of $1 \rightrightarrows B$ and is therefore a pullback of the diagonal $B \rightarrowtail B \times B$; hence it is complemented. □

Remark 4.5.8 In particular, it follows from 4.5.7 that *a topos which satisfies* (IC) *is necessarily Boolean.* This result (which was first proved by R. Diaconescu [281]) is sometimes viewed as a defect of topos theory by those constructivists who reject the impredicative power-set axiom, as discussed in Section D4.4: to them, the fact that the choiceness of 2 (equivalently, by 4.5.13 below, its well-orderability) implies the Law of Excluded Middle is further evidence that the existence of the power object $P2$ allows one 'too much freedom' to define subobjects of 2 which 'ought not to exist' in a truly constructive context. We prefer to see it as evidence of the remarkable power of the axiom of choice, and as a warning that it should not lightly be assumed in mathematical arguments.

What can we say about the class of choice objects in a general topos? With the exception of coproducts, it has good closure properties:

Lemma 4.5.9 *In any topos, the class of choice objects is closed under subobjects, quotients and finite products.*

Proof For subobjects, suppose we have $m \colon A' \rightarrowtail A$ where A is choice and $\phi \colon B \looparrowright A'$ is entire; then $m_\bullet \phi \colon B \looparrowright A$ is still entire, and any morphism contained in it must factor through m. For quotients, we have already observed that any quotient of a choice object appears as a subobject. We also observed that 1 is choice, in the proof of 4.5.7; for binary products, suppose A_1 and A_2 are choice and $\phi \colon B \looparrowright A_1 \times A_2$ is entire. Then we may first choose a morphism $f_1 \colon B \to A_1$ contained in $(\pi_1)_\bullet \phi$, then verify that $\phi \cap (\pi_1)^\bullet (f_1)_\bullet$ is still entire and choose a morphism $f_2 \colon B \to A_2$ contained in its composite with $(\pi_2)_\bullet$, and finally verify that $(f_1, f_2) \colon B \to A_1 \times A_2$ is contained in ϕ. □

Thus, in any Boolean topos, the full subcategory of choice objects is a sub-pretopos, in which every epimorphism splits. However, it is not a topos in general: it is well known that, in models of set theory without (AC), there exist sets A such that A is choice but PA is not. (On the other hand, it is interesting to

note that, in the topos $\mathbf{Cont}(G)$ of 4.5.2(*c*), the internally choice objects are exactly the uniformly continuous G-sets as defined in A2.1.7, at least provided we assume (AC) in $\mathbf{Set}$; so they do form a topos in this case.)

In $\mathbf{Set}$, Zermelo's Theorem tells us that the sets which admit choice functions are exactly those which can be well-ordered. The notion of well-ordering is definable in an arbitrary topos: an ordered object $(A, \leq)$ is said to be *well-ordered* if there exists a choice function $c\colon P^+A \to A$ such that the sentence asserting '$(\forall w\colon P^+A)(c(w)$ is the least element of $w)$' is satisfied. Clearly, a well-orderable object is a choice object; in the converse direction, the usual proof of Zermelo's Theorem uses the law of excluded middle, but is otherwise interpretable in the logic of a topos, so we know that any choice object in a Boolean topos admits a well-ordering.

Somewhat surprisingly, the same result holds in non-Boolean toposes. In order to prove this, we need a clearer picture of what choice objects look like in a non-Boolean topos. We shall need to use the notion of widespread subterminal object introduced in A1.6.11. Recall that, in any distributive lattice L, an element u is called widespread if the lattice $\{v \in L \mid v \geq u\}$ is complemented; and we say that a subterminal object U in a topos is (internally) widespread if $A \times U \rightarrowtail A$ is widespread in $\mathrm{Sub}(A)$ for all objects A – which is equivalent to saying that the sentence

$$(\forall p\colon \Omega)(p \vee (p \Rightarrow \ulcorner U \urcorner))$$

is valid in the internal logic of the topos, and to saying that the closed subtopos of $\mathcal{E}$ complementary to $\mathcal{E}/U$ (cf. A4.5.3) is Boolean. (Note that we have now ceased to distinguish notationally between a term t of type Ω and the corresponding formula $(* \in t)$ of the internal language, which we denoted $\bar{t}$ in Section D4.3.) More generally, we say a subobject $A' \rightarrowtail A$ is widespread in $\mathcal{E}$ if it is widespread as a subterminal object of $\mathcal{E}/A$. We may construct a subobject $W \rightarrowtail \Omega$ 'classifying' the class of widespread subobjects, namely the interpretation of

$$\{q\colon \Omega \mid (\forall p\colon \Omega)(p \vee (p \Rightarrow q))\} \; .$$

W is sometimes called the *Higgs object* of $\mathcal{E}$, in honour of D. Higgs who was the first to investigate it (in the context that we have already met in Section A1.6).

It is clear that, for any object B of $\mathcal{E}$, the forgetful functor $\Sigma_B\colon \mathcal{E}/B \to \mathcal{E}$ reflects the property of being choice, since it preserves entire relations. We shall also need

Lemma 4.5.10 *Let U be a subterminal object in a topos $\mathcal{E}$, and A an object which is a sheaf for the closed local operator $c(U)$. Then A is choice as an object of $\mathcal{E}$ iff it is choice as an object of $\mathbf{sh}_{c(U)}(\mathcal{E})$.*

Proof The assertion that A is a $c(U)$-sheaf says that $U^*(A) \cong 1_U$; so $U^*(P^+A) \cong P^+(1_U) \cong 1_U$, i.e. P^+A (as constructed in $\mathcal{E}$) is a $c(U)$-sheaf. It follows that the inclusion functor $\mathbf{sh}_{c(U)}(\mathcal{E}) \to \mathcal{E}$ preserves the construction of P^+, from which the result follows. □

Lemma 4.5.11 *Let A be a choice object of a topos $\mathcal{E}$, and suppose there exists a morphism $a: 1 \to A$. Then the classifying map of $a: 1 \rightarrowtail A$ factors through $W \rightarrowtail \Omega$.*

Proof Form the pushout

$$\begin{tikzcd} 1 \arrow[r, "a"] \arrow[d, "a"] & A \arrow[d] \\ A \arrow[r] & B \end{tikzcd} .$$

The two morphisms $(1_A \times a), (a \times 1_A): A \rightrightarrows A \times A$ induce a morphism $B \to A \times A$, which is monic by A1.4.3; so by 4.5.9 B is a choice object. It remains choice when we regard it as an object of $\mathcal{E}/A$ via the codiagonal map; but as an object of this topos, it lives in the closed subtopos $\mathbf{sh}_{c(a)}(\mathcal{E}/A)$, and in this subtopos it is simply the coproduct of two copies of the terminal object. So 4.5.10 says that $1 \amalg 1$ is a choice object in $\mathbf{sh}_{c(a)}(\mathcal{E}/A)$, and hence by 4.5.7 this topos is Boolean, i.e. a is internally widespread as a subterminal object of $\mathcal{E}/A$. □

Proposition 4.5.12 *In any topos, the Higgs object W is well-orderable; in particular, it is a choice object.*

Proof First we note that, if p and q are variables of type W, then the formulae $(p \vee (p \Rightarrow q))$ and $(q \vee (q \Rightarrow p))$ are valid, whence we have $(p \vee q \vee (p \Leftrightarrow q))$, or equivalently

$$(\forall p, q: W)((p = \top) \vee (q = \top) \vee (p = q)) .$$

(In other words, W is a torsor over the two-element monoid $\{1, e \mid e^2 = e\}$; cf. B3.2.4(c).) It follows easily that the binary relation on W defined by the term

$$\{\langle p, q\rangle : W \times W \mid ((p = \top) \vee (p = q))\}$$

is a total ordering (note that it is the opposite of the ordering induced by the usual ordering on Ω). We claim that it is a well-ordering.

To prove this, let $B \rightarrowtail W$ be a well-supported subobject. Form the pullback

$$\begin{tikzcd} U \arrow[r, tail] \arrow[d, tail] & 1 \arrow[d, tail, "\top"] \\ B \arrow[r, tail] & W \end{tikzcd} ;$$

then U is widespread as a subobject of B, so we can find a subobject $C \rightarrowtail B$ which is a complement for $B \times U$ in $\{B' \in \mathrm{Sub}(B) \mid U \leq B'\}$. We have

$\sigma C \cup \sigma(B \times U) = \sigma C \cup U = \sigma B = 1$; but $\sigma C \geq U$ since there exists a morphism $U \rightarrowtail C$, so $\sigma C = 1$. Now we have

$$(\forall q \colon B)((q = \top) \Rightarrow (\top \in \ulcorner B \urcorner)) ,$$

so from $(q = \top)$ we may deduce that $U = 1$ and hence that $C = U$. Thus

$$(\forall p \colon C)(\forall q \colon B)((q = \top) \Rightarrow (p = \top)) ,$$

and combining this with the property of W established in the first paragraph of the proof we have

$$(\forall p \colon C)(\forall q \colon B)((p = \top) \vee (p = q)) ,$$

that is, $(\forall p \colon C)(\forall q \colon B)(p \leq q)$. Combined with the fact that C is well-supported, this shows that $C \cong 1$ and that the image of $C \rightarrowtail B$ is the least element of B. Applying this argument to the generic well-supported subobject of W in the topos $\mathcal{E}/P^+W$, we deduce that the sentence of the internal language asserting the well-orderedness of W is satisfied. □

Theorem 4.5.13 *In any topos, the choice objects are exactly the well-orderable objects.*

Proof We have already observed that well-orderable objects are choice, so we have to prove the converse. Let A be a choice object in a topos $\mathcal{E}$; by working in the topos $\mathcal{E}/\sigma A$, where σA is the support of A, we may assume that A is well-supported, and hence that it has an element $a \colon 1 \to A$. By 4.5.11, the classifying map of a factors through $W \rightarrowtail \Omega$, so we may regard $(\chi_a \colon A \to W)$ as an object of $\mathcal{E}/W$. Also, since

$$\begin{tikzcd} 1 \arrow[r] \arrow[d, "a"'] & 1 \arrow[d, "\top"] \\ A \arrow[r, "\chi_a"] & W \end{tikzcd}$$

is a pullback, χ_a lives in the closed subtopos $\mathbf{sh}_{c(\top)}(\mathcal{E}/W)$ by A4.5.3(ii); and it is choice as an object of this topos, by 4.5.10. But this topos is Boolean because $\top$ is widespread in W; so by Zermelo's proof we may find a well-ordering of χ_a in this topos, and by an easy extension of 4.5.10 it remains a well-ordering of χ_a in $\mathcal{E}/W$.

We are thus reduced to proving that, given a morphism $f \colon A \to B$ in a topos $\mathcal{E}$, a well-ordering $\leq_B$ of B and a well-ordering $\leq_f$ of f in $\mathcal{E}/B$, the 'lexicographic' ordering of A given by

$$(x \leq_A y) \dashv\vdash ((fx \leq_B fy) \vee ((fx = fy) \wedge (x \leq_f y)))$$

is a well-ordering. But the well-known set-theoretic proof of this is constructive, and so valid in a topos: given an inhabited subobject $A' \rightarrowtail A$, we let

$B' = \exists_f(A') \rightarrowtail B$, find the $\leq_B$-least element $b\colon 1 \to B$ of B', and then find the least element of the pullback

$$\begin{tikzcd} A'' \arrow[r] \arrow[d, two heads] & A' \arrow[d, two heads] \\ 1 \arrow[r, "b"] & B' \end{tikzcd}$$

for the well-ordering which is the pullback of $\leq_f$ along b. □

Although, as we have seen, the axiom of choice in a topos is incompatible with non-Booleanness, there is a classically equivalent assertion, namely Zorn's Lemma, which can hold in non-Boolean toposes. It should be noted, however, that uses of Zorn's Lemma in mathematics are almost always followed immediately by an appeal to the law of excluded middle: having obtained a maximal element of some inductive poset, one assumes it does not have the property one is seeking and derives a contradiction by constructing a strictly larger element of the poset. Thus the validity of Zorn's Lemma in a non-Boolean topos is perhaps less useful than it appears at first sight – though we shall see one significant application of it in 4.6.15 below.

Given a poset $(A, \leq)$ in a topos, we say a subobject $A' \rightarrowtail A$ is a *chain* if it is totally ordered by (the restriction of) $\leq$, i.e. if it satisfies $(\forall x, y : A')$ $((x \leq y) \vee (y \leq x))$. And we say $(A, \leq)$ is *inductive* if it satisfies

$$(\forall w\colon PA)((w \text{ is a chain}) \Rightarrow (\exists x\colon A)(x \text{ is an upper bound for } w)) \, .$$

Proposition 4.5.14 *Assume* (AC) *holds in* **Set**, *and let* $\mathcal{E}$ *be a localic* **Set***-topos. If* $(A, \leq)$ *is an inductive poset in* $\mathcal{E}$, *then there exists* $a\colon 1 \to A$ *which is (internally) a maximal element of* A, *i.e. satisfies* $(\forall x\colon A)((a \leq x) \Rightarrow (a = x))$.

Proof Consider the (external!) poset C of chains in A, ordered by inclusion. It is clear that C is an inductive poset in **Set**, since the union of a chain of chains is a chain (and $\mathcal{E}$, being defined over **Set**, has arbitrary unions of subobjects). So it has an (externally) maximal element $A_0 \rightarrowtail A$, say. Since A is inductive, there exists $B \twoheadrightarrow 1$ in $\mathcal{E}$ and a morphism $a'\colon 1_B \to B^*A$ which is an upper bound for $B^*(A_0)$; we claim first that a' factors through $B^*(A_0)$, and is therefore the greatest element of this subobject. Since B can be covered by subterminal objects, it suffices to prove that, for every $f\colon U \to B$ with U subterminal, $f^*(a')$ factors through $U^*(A_0)$; but if U is subterminal, then $A_1 = A_0 \cup U \rightarrowtail A$ (where U maps to A by the transpose of $f^*(a')$) is easily seen to be a chain, and so must equal A_0 by maximality.

Now since greatest elements are unique when they exist, we see that the pullbacks of a' along the two projections $B \times B \rightrightarrows B$ must be equal; hence a' is itself the pullback along $B \to 1$ of a global element $a\colon 1 \to A$, which is the

greatest element of A_0. We claim that it is (internally) a maximal element of A. For, if we are given an object D and a morphism $b\colon 1_D \to D^*A$ satisfying $D^*a \leq b$, then it suffices as before to prove that b and D^*a coincide when pulled back along any $g\colon V \to D$ with V subterminal, and we may do this as before by constructing a chain $A_2 = A_0 \cup V$ containing A_0. □

Using 4.5.14, we may give a characterization of those **Set**-toposes which satisfy (AC):

Theorem 4.5.15 *Assume* (AC) *holds in* **Set**. *Then a* **Set**-*topos* ($\gamma\colon \mathcal{E} \to$ **Set**) *satisfies* (AC) *iff $\mathcal{E}$ is Boolean and γ is localic.*

Proof If $\mathcal{E}$ satisfies (AC), then it is Boolean by 4.5.8. We must show that the subterminal objects of $\mathcal{E}$ form a separating set: suppose given a parallel pair $f, g\colon A \rightrightarrows B$ with $f \neq g$. The equalizer $E \rightarrowtail A$ of f and g is not the whole of A, so its complement $E' \rightarrowtail A$ is nonzero; hence by (SS) we can find a nonzero subterminal object U and a morphism $h\colon U \to A$ factoring through E'. Now the equalizer of fh and gh is $h^*(E \rightarrowtail A) \cong (0 \rightarrowtail U)$, so $fh \neq gh$.

Conversely, suppose $\mathcal{E}$ is Boolean and γ is localic. Suppose given an epimorphism $f\colon A \twoheadrightarrow B$ in $\mathcal{E}$; then we may construct an internal poset $S \rightarrowtail P(A \times B)$ whose elements are (graphs of) partial sections of f, i.e. commutative diagrams

$$\begin{tikzcd} & A \arrow[d, "f"] \\ B' \arrow[ur, "s"] \arrow[r, tail] & B \end{tikzcd}$$

with the partial ordering given by $s \leq s'$ iff s' extends s. It is clear that S is inductive, so it has an (external) maximal element, corresponding to a partial section $s_0\colon B_0 \to A$. We must show that B_0 is the whole of B. Suppose not; then its complement $B_1 \rightarrowtail B$ is nonzero, and hence so is the pullback A_1 of A along $B_1 \rightarrowtail B$. Thus the coprojections $A_1 \rightrightarrows A_1 \amalg A_1$ are not equal, and so there exists a nonzero subterminal object U and a morphism $U \to A_1$. But now the composite $U \to A_1 \rightarrowtail A \to B$ is monic, and disjoint from $B_0 \rightarrowtail B$; so we may extend our section s_0 to one defined on $B_0 \amalg U$, contradicting the maximality of s_0. Thus we have shown that every epimorphism in $\mathcal{E}$ is split, i.e. $\mathcal{E}$ satisfies (AC). □

We have seen in the proof of 4.5.15 that if a topos is Boolean and satisfies (SS), then its subterminal objects form a separating set. The assumption of Booleanness is essential here: if $\mathcal{E}$ is the topos of sets equipped with an idempotent endomorphism, then $\mathcal{E}$ is 2-valued and its terminal object is projective, but not a separator.

Another application of (SS) allows us to 'lift' a weak form of the axiom of choice, namely (countable) dependent choices, from **Set** to at least some non-Boolean toposes. We recall that the *axiom of dependent choices* (DC) says that, if we are given a set A and a subset $R \subseteq A \times A$ which is entire as a relation $A \looparrowright A$ (i.e. such that $(\forall x : A)(\exists y : A)(\langle x, y \rangle \in R)$ is satisfied), then for any $a \in A$ there exists $f : \mathbb{N} \to A$ with $f(0) = a$ and $(\forall n : \mathbb{N})(\langle f(n), f(n+1) \rangle \in R)$. This axiom plays an important rôle in analysis; we shall meet it again in our discussion of the real numbers in a topos (see 4.7.12 below).

Lemma 4.5.16 *Suppose* $\mathcal{E} = \mathbf{Sh}(X)$ *is localic over* **Set** *and satisfies* (SS) *(and assume* (DC) *holds in* **Set***). Then* $\mathcal{E}$ *satisfies* (DC).

Proof Suppose given objects A and R satisfying the hypothesis of (DC) internally in $\mathcal{E}$. The key point is that, given an open $U \subseteq X$ and an element x of $A(U)$, we can find another element y of $A(U)$ such that $(\langle x, y \rangle \in \ulcorner R \urcorner)$ holds; for the object $[\![\{y : A \mid \langle x, y \rangle \in \ulcorner R \urcorner\}]\!]$ has support U by assumption. Thus, starting from any element a of $A(U)$, we can (using (DC) in **Set**) construct an infinite sequence of elements of $A(U)$ such that each is R-related to its successor; but since the natural number object in $\mathcal{E}$ is just the countable copower of 1, this is exactly what we need. □

Suggestions for further reading: Blass [123, 129], Diaconescu [281], Fourman & Ščedrov [365], Freyd [375], Penk [953].

D4.6 De Morgan's law and the Gleason cover

The assumption of Booleanness for a topos is in some ways an uncomfortably restrictive one, in that it eliminates most of the examples of toposes of greatest interest. However, there are various intermediate possibilities between the logic of an arbitrary topos and that of a Boolean topos; in fact there is a whole sequence of successively weaker conditions that we can add to the rules of intuitionistic propositional logic, corresponding to an increasing sequence of elements of the free Heyting algebra on one generator (the nth condition says that the nth element in the sequence is equal to $\top$). As they become weaker, these conditions become progressively less interesting (although, of course, there are progressively more toposes which satisfy them), but the first of them, which has come to be known as *De Morgan's law*, is of some interest as it arises naturally in both the logical and the geometrical aspects of topos theory. We devote the present section to studying it. (For information on the weaker conditions in the sequence, see [40, 41].)

In classical propositional logic, the name 'De Morgan's laws' is given to the sequents

$$(\neg(p \vee q) \dashv\vdash (\neg p \wedge \neg q)) \quad \text{and} \quad (\neg(p \wedge q) \dashv\vdash (\neg p \vee \neg q))$$

which express the fact that negation is a lattice anti-isomorphism in a Boolean algebra. The first of these two laws is in fact valid in any Heyting algebra: since $\neg(-) = ((-) \Rightarrow \bot)$ is self-adjoint on the right, it converts joins to meets. But the second one fails in general: the right-to-left entailment is still valid, but the left-to-right one is not.

Lemma 4.6.1 *Let H be a Heyting algebra. The following conditions are equivalent:*

(i) *For all $p, q \in H$, $\neg(p \wedge q) = (\neg p \vee \neg q)$.*

(ii) *For all $p \in H$, $(\neg p \vee \neg\neg p) = \top$.*

(iii) *Every $\neg\neg$-fixed element of H is complemented.*

(iv) *The $\neg\neg$-fixed elements of H form a sublattice.*

(v) *$\neg\neg\colon H \to H$ is a lattice homomorphism.*

Proof (i) $\Rightarrow$ (ii): Substitute $\neg p$ for q in (i). Since $(p \wedge \neg p) = \bot$, we deduce the identity of (ii).

(ii) $\Rightarrow$ (iii): (ii) says that $\neg\neg p$ has a complement, namely $\neg p$.

(iii) $\Rightarrow$ (iv): A complemented element is always $\neg\neg$-fixed, so (iii) implies that the $\neg\neg$-fixed elements of H coincide with the complemented elements; but the latter always form a sublattice.

(iv) $\Rightarrow$ (v): $\neg\neg$ always preserves $\bot$, $\top$ and binary meets; and it preserves binary joins as a map from H to the lattice $H_{\neg\neg}$ of $\neg\neg$-fixed elements, since it is left adjoint to the inclusion. So (iv) implies that it preserves binary joins as a map $H \to H$.

(v) $\Rightarrow$ (i): If (v) holds, then we have

$$\begin{aligned}(\neg p \vee \neg q) &= (\neg\neg\neg p \vee \neg\neg\neg q)\\ &= \neg\neg(\neg p \vee \neg q)\\ &= \neg(\neg\neg p \wedge \neg\neg q)\\ &= \neg\neg\neg(p \wedge q)\\ &= \neg(p \wedge q)\end{aligned}$$

using the identity $\neg = \neg\neg\neg$ and the fact that the first De Morgan law always holds. □

A Heyting algebra satisfying the equivalent conditions of 4.6.1 is sometimes called a *Stone algebra.* We shall say that a topos $\mathcal{E}$ satisfies *De Morgan's law* if its subobject classifier is an internal Stone algebra in $\mathcal{E}$; equivalently, if for every subobject $A' \rightarrowtail A$ in $\mathcal{E}$ we have $\neg A' \vee \neg\neg A' \cong A$ in $\mathrm{Sub}(A)$. There are many different ways of expressing this condition: the next proposition lists a selection of them.

Proposition 4.6.2 *The following conditions on a topos $\mathcal{E}$ are equivalent:*

(i) *$\mathcal{E}$ satisfies De Morgan's law.*

(ii) *Every $\neg\neg$-closed subobject in $\mathcal{E}$ has a complement.*

(iii) *The subobject* $\perp : 1 \rightarrowtail \Omega$ *has a complement.*

(iv) *An object of* $\mathcal{E}$ *is separated for the local operator* $\neg\neg$ *(cf.* A4.5.9*) iff it is decidable.*

(v) *Every* $\neg\neg$*-sheaf in* $\mathcal{E}$ *is decidable.*

(vi) $\Omega_{\neg\neg}$ *is decidable in* $\mathcal{E}$.

(vii) *The canonical monomorphism* $(\top, \perp) : 2 = 1 \amalg 1 \rightarrowtail \Omega_{\neg\neg}$ *is an isomorphism in* $\mathcal{E}$.

(viii) 2 *is an internally complete poset in* $\mathcal{E}$.

(ix) 2 *is an injective object of* $\mathcal{E}$.

(x) *The class of injective objects of* $\mathcal{E}$ *is closed under binary coproducts.*

(xi) 2 *is a retract of* Ω *in* $\mathcal{E}$.

(xii) 2 *is a* $\neg\neg$*-sheaf in* $\mathcal{E}$.

(xiii) *The subcategory* $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ *is closed under finite coproducts.*

Proof (i) $\Leftrightarrow$ (ii) is immediate from 4.6.1; and (ii) $\Leftrightarrow$ (iii) because a subobject is $\neg\neg$-closed iff it occurs as the negation of something, iff it is a pullback of $\perp$.

(ii) $\Rightarrow$ (iv): By A4.3.6(a), an object A is separated for a local operator j iff its diagonal is j-closed.

(iv) $\Rightarrow$ (v) is trivial; and (v) $\Rightarrow$ (vi) since $\Omega_{\neg\neg}$ is a $\neg\neg$-sheaf by A4.3.7.

(vi) $\Rightarrow$ (vii): If (v) holds, then $\top : 1 \to \Omega_{\neg\neg}$ has a complement, which must be $\perp : 1 \to \Omega_{\neg\neg}$ since the latter is the largest subobject of Ω disjoint from $\top$.

(vii) $\Rightarrow$ (ii) since 2 classifies complemented subobjects in $\mathcal{E}$, and $\Omega_{\neg\neg}$ classifies $\neg\neg$-closed subobjects.

(vii) $\Rightarrow$ (viii) since $\Omega_{\neg\neg}$, being reflective in Ω, is always an internally complete poset in a topos.

(viii) $\Rightarrow$ (ix): The underlying object of an internally complete poset is always injective, since it is a retract of a power object by B2.3.9(iv).

(ix) $\Leftrightarrow$ (x): Suppose (ix) holds: let A and B be injective objects, and suppose we are given a monomorphism $m : A \amalg B \rightarrowtail C$. By injectivity of 2, we may find a commutative triangle

$$\begin{tikzcd} A \amalg B \arrow[r, tail, "m"] \arrow[d] & C \arrow[dl] \\ 1 \amalg 1 & \end{tikzcd}$$

and hence a coproduct decomposition $C \cong C_1 \amalg C_2$ such that m restricts to monomorphisms $A \rightarrowtail C_1$ and $B \rightarrowtail C_2$. Then we may split these monomorphisms to obtain a splitting for m. Since monomorphisms are stable under pushout in a topos, this is sufficient. The converse is immediate since 1 is always injective.

(ix) $\Leftrightarrow$ (xi): 2 is always a subobject of Ω; and Ω is always injective by A2.2.6.

(xi) $\Rightarrow$ (xii): Since 2 is always decidable (cf. A1.4.15), it is $\neg\neg$-separated; injectivity forces it to be a sheaf for any local operator for which it is separated.

(xii) $\Leftrightarrow$ (xiii): Suppose (xii) holds; let A_1 and A_2 be $\neg\neg$-sheaves, and suppose we are given a diagram

$$\begin{tikzcd} B \arrow[r, "f"] \arrow[d, tail, "m"] & A_1 \amalg A_2 \\ C & \end{tikzcd}$$

with m $\neg\neg$-dense. Then the composite $B \to A_1 \amalg A_2 \to 1 \amalg 1$ has a unique extension along m; so we obtain coproduct decompositions $C \cong C_1 \amalg C_2$ and $B \cong B_1 \amalg B_2$ such that the morphisms f and m similarly decompose as $f_1 \amalg f_2$ and $m_1 \amalg m_2$. Then we may extend f_1 uniquely along m_1 and f_2 uniquely along m_2 (since m_1 and m_2, being pullbacks of m along the coprojections, are $\neg\neg$-dense), and verify that these combine to yield the unique extension of f along m. Thus a binary coproduct of $\neg\neg$-sheaves is a $\neg\neg$-sheaf; but the empty coproduct 0 is always a $\neg\neg$-sheaf. The converse is again immediate since 1 is always a sheaf.

(xii) $\Rightarrow$ (vii): Since $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ is Boolean, its subobject classifier $\Omega_{\neg\neg}$ is the coproduct of two copies of 1 in it; in other words, it is the associated $\neg\neg$-sheaf of 2, the unit of the adjunction being the morphism in (vii). So if 2 is itself a $\neg\neg$-sheaf, this morphism is an isomorphism. $\square$

For yet more conditions equivalent to De Morgan's law in a topos, see 4.6.14 and 4.7.11 below.

Examples 4.6.3 (*a*) Let $\mathcal{C}$ be a small category. We recall that the subobject classifier in $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$ is defined by

$$\Omega(A) = \{\text{sieves on } A \text{ in } \mathcal{C}\}\ ,$$

and $\perp\colon 1 \to \Omega$ is the natural transformation which picks out the empty sieve on each object of $\mathcal{C}$. So 4.6.2(ii) is equivalent to the statement that the collection of nonempty sieves forms a subfunctor Ω' of Ω, i.e. that the pullback of a nonempty sieve along an arbitrary morphism is nonempty. But, as we saw in A2.1.11(*h*), this in turn is equivalent to saying that $\mathcal{C}$ satisfies the 'right Ore condition' that any pair of morphisms with common codomain can be embedded in a commutative square. Note in particular that this condition holds if $\mathcal{C}$ has pullbacks; thus, for example, we may conclude from 3.1.1 that *the clasifying topos (over* **Set***) of any cartesian theory satisfies De Morgan's law.* (This may be contrasted with the very special nature of those coherent theories whose classifying toposes are Boolean; cf. 3.4.6. And see also C4.3.2, for a class of Grothendieck toposes which always satisfy De Morgan's law but are almost never Boolean.)

(*b*) For a locale X, we saw in C1.5.10 that $\mathbf{Sh}(X)$ satisfies De Morgan's law iff X is *extremally disconnected*, i.e. iff the closure of every open sublocale of X is open.

(*c*) A particular case of (*b*) which will be of importance later in this section: if B is a Boolean algebra (regarded as a coherent category) and C is its coherent coverage (cf. A2.1.11(*b*)), then $\mathbf{Sh}(B, C)$ *satisfies De Morgan's law iff* B *is complete*. Although, as just stated, this looks rather different from (*b*), it is a special case because C2.2.4(*b*) tells us that $\mathbf{Sh}(B, C)$ is equivalent to $\mathbf{Sh}(X)$ where X is the Stone space of B (see [520]), that is the locale corresponding to the frame $I(B)$ of ideals of B. It is well known that this locale is extremally disconnected iff B is complete, but we can give a direct proof as follows. Let J be an ideal of B; then an element b of B belongs to its negation $\neg J$ (in $I(B)$) iff $b \wedge c = 0$ for all $c \in J$, iff its complement $\neg b$ is an upper bound for J. So $\neg J$ is a principal ideal, for all ideals J of B, iff every ideal has a least upper bound, iff B is complete. But to say that the negation of every ideal is principal implies that $I(B)$ is an (external) Stone algebra, since the principal ideals form a sublattice. Conversely, if B is complete, we may 'relativize' the above argument to every principal ideal of B, to deduce that the subobject classifier of $\mathbf{Sh}(B, C)$ is an internal Stone algebra.

It will be seen that the key condition in 4.6.2 is (vii), which says that 2 coincides with $\Omega_{\neg\neg}$; most of the other conditions say either that 2 (resp. the class of all decidable objects) has some property which is known to hold for $\Omega_{\neg\neg}$ (resp. the class of all $\neg\neg$-sheaves) in any topos, or that $\Omega_{\neg\neg}$ (resp. the class of all $\neg\neg$-sheaves) has some property which is known to hold for 2 (resp. the class of all decidable objects) in any topos. One further condition of this kind, which looks as if it 'ought to be' equivalent to those of 4.6.2 but turns out to be slightly weaker, is contained in the following lemma.

Lemma 4.6.4 *The following conditions on a topos $\mathcal{E}$ are equivalent, and are implied by those of* 4.6.2:

(i) 2^A *is decidable for any object A of $\mathcal{E}$.*

(ii) *The decidable objects form an exponential ideal in $\mathcal{E}$.*

Proof Condition (ii) is implied by 4.6.2(iv), since the separated objects for any local operator form an exponential ideal by A4.4.3(ii). It is clear that (ii) implies (i); for the converse, suppose B is decidable. Then we have a pullback diagram of the form

$$\begin{tikzcd} B \arrow[r] \arrow[d, tail, "\Delta"] & 1 \arrow[d, tail] \\ B \times B \arrow[r] & 2 \end{tikzcd}$$

which is preserved by the functor $(-)^A$ for any A. If 2^A is decidable, then any morphism $1 \cong 1^A \to 2^A$ is a complemented subobject; hence the diagonal $\Delta\colon B^A \rightarrowtail B^A \times B^A$ is complemented. □

Example 4.6.5 Let $\mathcal{C}$ be a small category satisfying the condition that, for every $f\colon A \to B$ in $\mathcal{C}$, there exist $g\colon B \to A$ and $h\colon B \to C$ such that $h = hfg$. Then the topos $[\mathcal{C}, \mathbf{Set}]$ satisfies the conditions of 4.6.4. For we recall from A1.4.16 that a functor $F\colon \mathcal{C} \to \mathbf{Set}$ is decidable in $[\mathcal{C}, \mathbf{Set}]$ iff $F(f)$ is injective for all f, and from A1.5.5 that the exponential G^F in $[\mathcal{C}, \mathbf{Set}]$ is defined by setting $G^F(A)$ to be the set of natural transformations $\mathcal{C}(A, -) \times F \to G$. So suppose G is decidable, and let α, β be two such natural transformations which have the same composite with $\mathcal{C}(k, -) \times 1_F$ for some $k\colon A \to D$; in other words, they agree on all pairs of the form (f, x) where f factors through k. Now consider $\alpha_B(f, x)$ for an arbitrary $f\colon A \to B$ and $x \in F(B)$; choosing g and h as in the condition above, we have

$$\begin{aligned} G(h)(\alpha_B(f,x)) &= \alpha_C(hf, F(f)(x)) \\ &= \alpha_C(hf, F(hfg)(x)) \\ &= G(hf)\alpha_A(1_A, F(g)(x)) \,. \end{aligned}$$

But since $G(k)$ is injective and $\alpha_D(k, F(kg)(x)) = \beta_D(k, F(kg)(x))$, we have

$$\alpha_A(1_A, F(g)(x)) = \beta_A(1_A, F(g)(x)) \;;$$

hence $G(h)(\alpha_B(f,x)) = G(h)(\beta_B(f,x))$, and since $G(h)$ is injective we have $\alpha_B(f,x) = \beta_B(f,x)$. Thus $\alpha = \beta$; so we have shown that $G^F(k)$ is injective for an arbitrary morphism k of $\mathcal{C}$, i.e. G^F is decidable.

Now take $\mathcal{C}$ to be the three-element monoid $\{1, a, b\}$ with $a^2 = ba = a$ and $b^2 = ab = b$. Since the morphisms of $\mathcal{C}$ are all idempotent, it is easy to see that $\mathcal{C}$ satisfies the condition above; but it does not satisfy the left Ore condition. So the topos $[\mathcal{C}, \mathbf{Set}]$ satisfies the conditions of 4.6.4, but not those of 4.6.2.

We may also mention a stronger condition than De Morgan's law: we say that a topos satisfies the *strong De Morgan's law* if its subobject classifier is internally totally ordered, i.e. if the sentence

$$(\forall p, q\colon \Omega)(((p \Rightarrow q) \vee (q \Rightarrow p)) = \top)$$

is satisfied. It is easy to see that this implies De Morgan's law (in the form given in 4.6.1(ii)), by substituting $\neg p$ for q, but the converse is false. For a functor category $[\mathcal{C}^{\mathrm{op}}, \mathbf{Set}]$, it is easily seen to be equivalent to the condition that every slice category $\mathcal{C}/A$ is strongly connected (i.e., given any two morphisms f, g of $\mathcal{C}$ with common codomain, either f factors through g or g factors through f) – which is clearly stronger than the Ore condition of 4.6.3(a). However, we have

the following result:

Lemma 4.6.6 *For a topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ satisfies the strong De Morgan's law.*

(ii) *For every object A of $\mathcal{E}$ and every local operator j in $\mathcal{E}/A$, $\mathbf{sh}_j(\mathcal{E}/A)$ satisfies the strong De Morgan's law.*

(iii) *For every object A of $\mathcal{E}$ and every closed local operator c in $\mathcal{E}/A$, $\mathbf{sh}_c(\mathcal{E}/A)$ satisfies De Morgan's law.*

Proof (i)$\Rightarrow$(ii): Since the strong De Morgan's law is expressible by a sentence in the internal language, it is clearly inherited by slice categories; so we need only show that it is inherited by sheaf subtoposes. But, if j is a local operator in $\mathcal{E}$, then Ω_j is a sub-poset of Ω and hence totally ordered in $\mathcal{E}$; and the associated sheaf functor preserves total orderings, so it is totally ordered in $\mathbf{sh}_j(\mathcal{E})$.

(ii)$\Rightarrow$(iii) is trivial.

(iii)$\Rightarrow$(i): If p and q are variables of type Ω, then $(p \wedge q = \bot)$ is easily seen to imply $((p \Rightarrow q) \vee (q \Rightarrow p)) = (\neg p \vee \neg q) \geq (\neg p \vee \neg\neg p)$; so the strong form of De Morgan's law follows from the ordinary version in this particular case. Thus, to verify that the strong form holds in $\mathcal{E}$, it is sufficient to know that the ordinary form holds in the closed subtopos of $\mathcal{E}/\Omega \times \Omega$ complementary to the subterminal object $(\top, \top): 1 \rightarrowtail \Omega \times \Omega$ which is classified by the meet of the two projection maps $\Omega \times \Omega \rightrightarrows \Omega$. □

Using this result, it is easy to verify that a localic topos $\mathbf{Sh}(X)$ satisfies the strong De Morgan's law iff every closed sublocale of X is extremally disconnected. (It can be shown directly that this is equivalent to saying that every sublocale of X is extremally disconnected, using the fact that any sublocale is an intersection of sublocales which are the union of an open and a closed sublocale; cf. C1.1.17.)

Because of the presence of negation in De Morgan's law, we cannot expect its validity to be preserved or reflected along geometric morphisms, in general. However, there are two particular cases of interest:

Lemma 4.6.7

(i) *If $f: \mathcal{F} \to \mathcal{E}$ is an open surjection, and $\mathcal{F}$ satisfies De Morgan's law, then so does $\mathcal{E}$.*

(ii) *If $\mathcal{E}$ is a retract in $\mathfrak{Top}$ of a topos satisfying De Morgan's law, then $\mathcal{E}$ satisfies De Morgan's law.*

Proof (i) is immediate from the fact that f^* is a conservative Heyting functor.

(ii) Let $r: \mathcal{F} \to \mathcal{E}$, $i: \mathcal{E} \to \mathcal{F}$ be geometric morphisms with $ri \cong 1_{\mathcal{E}}$. Given an arbitrary subobject $A' \rightarrowtail A$ in $\mathcal{E}$, we have $r^*(\neg A') \leq \neg(r^*(A'))$ in $\mathrm{Sub}(r^*(A))$, since $r^*(\neg A')$ is disjoint from $r^*(A')$. The functor i^* preserves this inequality, but we also have $i^*(\neg(r^*(A'))) \leq \neg(i^*r^*(A')) \cong \neg A'$ in $\mathrm{Sub}(A)$, for a similar reason, and hence $\neg A' \cong i^*(\neg(r^*(A')))$. Since i^* preserves complemented subobjects, it follows that $\neg A'$ is complemented; so $\mathcal{E}$ satisfies De Morgan's law. □

We remark that the analogues of both parts of 4.6.7 hold for the strong De Morgan's law of 4.6.6: they may be proved by similar arguments, or alternatively deduced from 4.6.6(iii).

We saw in A4.5.23 that, for any topos $\mathcal{E}$, we can find a (geometric) surjection $f: \mathcal{F} \to \mathcal{E}$ with $\mathcal{F}$ Boolean. Indeed, we have a canonical construction of such a topos $\mathcal{F}$; but the morphism f does not have any particularly good properties (apart from being localic). If we make the more modest demand that the 'covering topos' $\mathcal{F}$ should satisfy De Morgan's law, then we can do rather better: by 4.6.7(i), we cannot hope for the covering morphism to be an open surjection, but it turns out that we can take it to be a proper surjection (indeed, tidy in the sense of C3.4.2), and hence a descent morphism.

Before stating the next result, we need a new definition: we say a localic surjection $f: \mathcal{F} \to \mathcal{E}$ is *minimal* if the only closed subtopos of $\mathcal{F}$ which maps surjectively to $\mathcal{E}$ is $\mathcal{F}$ itself. If f is a closed map in the sense of C3.2.1, so that the image of a closed subtopos $\mathbf{sh}_{c(U)}(\mathcal{F})$ (where $U \in \mathrm{Sub}_{\mathcal{F}}(1)$) is simply given by $\mathbf{sh}_{c(f_*(U))}(\mathcal{E})$, this is equivalent to saying that $f_*: \mathrm{Sub}_{\mathcal{F}}(1) \to \mathrm{Sub}_{\mathcal{E}}(1)$ reflects, as well as preserving, the bottom element.

Theorem 4.6.8 *For any topos $\mathcal{E}$, there exists a topos $\gamma\mathcal{E}$ satisfying De Morgan's law and a surjective geometric morphism $f: \gamma\mathcal{E} \to \mathcal{E}$ which is localic, proper (in the sense of* C3.2.12*), separated (in the sense of the same reference) and minimal in the sense defined above. Moreover, $\gamma\mathcal{E}$ is characterized up to equivalence in $\mathfrak{Top}/\mathcal{E}$ by these properties.*

Proof We define $\gamma\mathcal{E}$ to be the topos of sheaves on the internal site in $\mathcal{E}$ obtained by equipping the internal Boolean algebra $\Omega_{\neg\neg}$ with its coherent coverage; equivalently, it is the topos of sheaves on the internal frame $I(\Omega_{\neg\neg})$ of ideals of $\Omega_{\neg\neg}$. Since $\Omega_{\neg\neg}$ is a complete Boolean algebra, it follows from 4.6.3(c) (which did not make any appeal to the classical logic of **Set**) that this topos indeed satisfies De Morgan's law. Moreover, $I(\Omega_{\neg\neg})$ is a coherent frame, so it is compact (indeed, strongly compact, by C3.4.1(c)); it is also regular, since the principal ideals are complemented and form a basis for it, and so Hausdorff by C1.2.17(i). Hence $\gamma\mathcal{E} \to \mathcal{E}$ is proper and separated. For minimality, we note that if J is an ideal of $\Omega_{\neg\neg}$ (that is, a subterminal object of $\gamma\mathcal{E}$), then $f_*(J)$ is simply the truth-value $[\![\top \in J]\!]$; so $f_*(J) = \bot$ is equivalent to $\neg(\top \in J)$, and hence to $J = \{\bot\}$ (since any ideal must contain $\bot$, and $\{\bot\}$ is the pseudocomplement of $\{\top\}$ in $\Omega_{\neg\neg}$).

For the converse, suppose $f: \mathcal{F} \to \mathcal{E}$ is a geometric morphism with the stated properties. Since f is localic, it corresponds to an internal locale X in $\mathcal{E}$; we must show that $\mathcal{O}(X)$ is isomorphic to $I(\Omega_{\neg\neg})$. Since f is proper and separated, X is a compact Hausdorff locale, and hence regular by C3.2.10(ii). Moreover, it is extremally disconnected since $\mathcal{F}$ satisfies De Morgan's law; so, from the information that every $U \in \mathcal{O}(X)$ is a union of opens whose closures are contained in U, we deduce that every $U \in \mathcal{O}(X)$ is a union of complemented (and hence compact) elements of $\mathcal{O}(X)$. Thus $\mathcal{O}(X) \cong I(B)$, where B is the sub-semilattice

of compact elements of $\mathcal{O}(X)$; and since the compact elements are exactly the complemented ones by C3.2.10(i), B is actually a Boolean algebra – and it is complete, by 4.6.3(c). Finally, the minimal surjectivity of f tells us, as before, that the Heyting pseudocomplement of the singleton $\{1_B\}$ in Sub(B) must be $\{0_B\}$ – equivalently, since complementation is an isomorphism $B \to B$, $\{1_B\}$ is the pseudocomplement of $\{0_B\}$. So the classifying map $\chi\colon B \to \Omega$ of $\{1_B\}$ factors through $\Omega_{\neg\neg}$; and since we have

$$(\forall b\colon B)\neg\neg((b = 1_B) \vee (b = 0_B))$$

(as the union of a subobject and its pseudocomplement is always $\neg\neg$-dense), it is straightforward to verify that the induced factorization $B \to \Omega_{\neg\neg}$ is an isomorphism of Boolean algebras. □

The topos $\gamma\mathcal{E}$ is called the *Gleason cover* of $\mathcal{E}$ in honour of A. M. Gleason [409] who was the first to investigate the corresponding problem of covering an arbitrary space by an extremally disconnected one. The construction for toposes was first given in [512]. Note that the morphism $\gamma\mathcal{E} \to \mathcal{E}$ is entire in the sense of B4.4.21(a); indeed, we could have replaced 'proper and separated' by 'entire' in the characterization above.

Note also that if $\mathcal{E}$ itself satisfies De Morgan's law, then $\gamma\mathcal{E} \to \mathcal{E}$ is an equivalence; this may of course be deduced from the characterization of it in 4.6.8, but also follows directly from the definition since if $\mathcal{E}$ satisfies De Morgan's law we have $I(\Omega_{\neg\neg}) \cong I(2) \cong \Omega$. This suggests that the construction $\mathcal{E} \mapsto \gamma\mathcal{E}$ should be a coreflection of some kind; indeed it is, although the category on which it is functorial is far from being the whole of $\mathfrak{Top}$.

Lemma 4.6.9 *For a geometric morphism $f\colon \mathcal{F} \to \mathcal{E}$, the following are equivalent:*

(i) *f restricts to a geometric morphism* $\mathbf{sh}_{\neg\neg}(\mathcal{F}) \to \mathbf{sh}_{\neg\neg}(\mathcal{E})$.

(ii) *f^* preserves $\neg\neg$-dense monomorphisms.*

(iii) *For any monomorphism $A' \rightarrowtail A$ in $\mathcal{E}$, we have $\neg\neg f^*(A') \cong \neg f^*(\neg A')$ in* $\mathrm{Sub}_{\mathcal{F}}(f^*(A))$.

Proof (i)$\Leftrightarrow$(ii) is a particular case of A4.3.12.

(ii)$\Rightarrow$(iii): If (ii) holds, then $f^*(A') \vee f^*(\neg A') \rightarrowtail f^*(A)$ must be $\neg\neg$-dense for any $A' \rightarrowtail A$; pulling back along $\neg f^*(A') \rightarrowtail f^*(A)$ we deduce that $f^*(\neg A') \rightarrowtail \neg f^*(A')$ is $\neg\neg$-dense (note that the former is always contained in the latter, since it is disjoint from $f^*(A')$), and taking pseudocomplements we obtain (iii).

(iii)$\Rightarrow$(ii): If (iii) holds, then on applying it to a $\neg\neg$-dense subobject $A' \rightarrowtail A$ we deduce that $\neg\neg(f^*A') \cong \neg f^*(0) \cong f^*(A)$. □

Geometric morphisms satisfying the conditions of 4.6.9 are commonly called *skeletal*. They include all open maps in the sense of C3.1.7 (since the inverse

image of an open map commutes with negation), but also all dense inclusions since if $\mathcal{F}$ is a dense subtopos of $\mathcal{E}$ then $\mathbf{sh}_{\neg\neg}(\mathcal{F})$ coincides with $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ (cf. A4.5.21). In fact we can characterize skeletal inclusions as follows:

Lemma 4.6.10 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be an inclusion, corresponding to a local operator j on $\mathcal{E}$. Then f is skeletal iff the exterior of j (that is, the open complement of the closure of j, cf.* A4.5.19(i)*) is a $\neg\neg$-closed subterminal object of $\mathcal{E}$.*

Proof First suppose f is skeletal. From the proof of A4.5.19(i), we recall that $\mathrm{ext}(j)$ is the image of the initial object 0 of $\mathcal{F}$ under f_*; but 0 is a $\neg\neg$-sheaf in $\mathcal{F}$, so $\mathrm{ext}(j)$ must be a $\neg\neg$-sheaf in $\mathcal{E}$. Equivalently, by A4.3.8, it is a $\neg\neg$-closed subobject of 1.

Conversely, suppose the condition is satisfied, and consider the pullback

$$\begin{tikzcd} \mathcal{F}' \arrow[r] \arrow[d] & \mathbf{sh}_{\neg\neg}(\mathcal{E}) \arrow[d] \\ \mathcal{F} \simeq \mathbf{sh}_j(\mathcal{E}) \arrow[r, "f"] & \mathcal{E} \end{tikzcd} .$$

By A4.5.16, $\mathcal{F}'$ is simply the intersection of the two subcategories $\mathbf{sh}_j(\mathcal{E})$ and $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ of $\mathcal{E}$; in particular it contains the object $\mathrm{ext}(j)$, since the given condition implies that this is a $\neg\neg$-sheaf as well as a j-sheaf. So $\mathcal{F}'$ contains the initial object of $\mathcal{F}$, and is therefore dense; hence it contains the subtopos $\mathbf{sh}_{\neg\neg}(\mathcal{F})$, and so f is skeletal. (In fact $\mathcal{F}'$ must coincide with $\mathbf{sh}_{\neg\neg}(\mathcal{F})$ in this case, since it is a subtopos of $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ and therefore Boolean by A4.5.22.) □

Corollary 4.6.11 *The class of skeletal inclusions is the smallest class Σ such that*

(i) *open inclusions are in Σ, and*

(ii) *given a composable pair of inclusions*

$$\mathcal{G} \xrightarrow{g} \mathcal{F} \xrightarrow{f} \mathcal{E}$$

where g is dense, we have $f \in \Sigma$ iff $fg \in \Sigma$.

Proof The class of skeletal inclusions certainly has these closure properties (the second one because, if g is dense, then $\mathbf{sh}_{\neg\neg}(\mathcal{F})$ and $\mathbf{sh}_{\neg\neg}(\mathcal{G})$ coincide, as we recalled earlier). Conversely, suppose $f\colon \mathbf{sh}_j(\mathcal{E}) \to \mathcal{E}$ is skeletal, and let U be the pseudo-complement of $\mathrm{ext}(j)$ in $\mathrm{Sub}_{\mathcal{E}}(1)$. Then $\mathrm{ext}(o(U)) = \neg U = \mathrm{ext}(j)$ by 4.6.10, so the open subtopos $\mathbf{sh}_{o(U)}(\mathcal{E})$ and $\mathbf{sh}_j(\mathcal{E})$ are both dense in the closed subtopos $\mathbf{sh}_{c(\mathrm{ext}(j))}(\mathcal{E})$. Hence $f \in \Sigma$. □

For more information about the composition properties of skeletal morphisms, see [519].

It is easy to see that $\gamma\mathcal{E} \to \mathcal{E}$ is always skeletal; for we can represent $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ as the topos of sheaves for the canonical coverage C on $\Omega_{\neg\neg}$, and the inclusion $\mathbf{Sh}_{\mathcal{E}}(\Omega_{\neg\neg}, C) \to \mathbf{Sh}_{\mathcal{E}}(\Omega_{\neg\neg}, P)$ (where P is the coherent coverage) is clearly dense. Indeed, since 4.6.2(vii) says that the inclusion $\mathbf{sh}_{\neg\neg}(\mathcal{E}) \simeq \mathbf{sh}_{\neg\neg}(\gamma\mathcal{E}) \to \gamma\mathcal{E}$ is pure in the sense of C3.4.12, we could alternatively have constructed $\gamma\mathcal{E}$ as (the middle term of) the pure–entire factorization of $\mathbf{sh}_{\neg\neg}(\mathcal{E}) \to \mathcal{E}$. From the latter observation, and the functoriality of the pure–entire factorization, we immediately deduce

Proposition 4.6.12 *Let $\mathfrak{Top}_{sk}$ denote the 2-category of toposes, skeletal geometric morphisms and geometric transformations between them, and $\mathfrak{DMTop}_{sk}$ the full sub-2-category whose objects are toposes satisfying De Morgan's law. Then $\mathcal{E} \mapsto \gamma\mathcal{E}$ is a functor $\mathfrak{Top}_{sk} \to \mathfrak{DMTop}_{sk}$, right adjoint to the inclusion.*

Proof The '1-dimensional' part of the proof is contained in the discussion above; to complete it, we have only to discuss the effect of γ on geometric transformations. For this we use an idea that we have exploited before in B3.2.14, namely that the Sierpiński topos over $\mathcal{F}$ (i.e. the diagram category $[\mathbf{2}, \mathcal{F}]$) may be regarded as the tensor $\mathbf{2} \otimes \mathcal{F}$ in $\mathfrak{Top}$. Now the subobject classifier in $[\mathbf{2}, \mathcal{F}]$ may, by the construction of A2.1.12, be identified with the diagram $(\vee \colon \Omega_1 \to \Omega)$, where Ω is the subobject classifier of $\mathcal{F}$ and Ω_1 is, as usual, its order-relation; and an easy calculation shows that $\Omega_{\neg\neg}$ in $[\mathbf{2}, \mathcal{F}]$ is simply the identity morphism $(\Omega_{\neg\neg} \to \Omega_{\neg\neg})$ (where the inclusion into $\Omega_{[\mathbf{2},\mathcal{F}]}$ maps the first $\Omega_{\neg\neg}$ into Ω_1 via the diagonal). In particular, we see that if $\mathcal{F}$ satisfies De Morgan's law, so that $\Omega_{\neg\neg}$ coincides with 2 in $\mathcal{F}$, then so does $[\mathbf{2}, \mathcal{F}]$. Also, for any $\mathcal{F}$, we have $\gamma[\mathbf{2}, \mathcal{F}] \cong [\mathbf{2}, \gamma\mathcal{F}]$; the topos $\mathbf{sh}_{\neg\neg}([\mathbf{2}, \mathcal{F}])$ does not coincide with $[\mathbf{2}, \mathbf{sh}_{\neg\neg}(\mathcal{F})]$ (since the latter is not Boolean!) but with $\mathbf{sh}_{\neg\neg}(\mathcal{F})$, mapping into $[\mathbf{2}, \mathcal{F}]$ via the open inclusion $\mathcal{F} \to [\mathbf{2}, \mathcal{F}]$. Hence the projection $[\mathbf{2}, \mathcal{F}] \to \mathcal{F}$, and the open inclusion $\mathcal{F} \to [\mathbf{2}, \mathcal{F}]$, are both skeletal; the closed inclusion is not skeletal (so that $[\mathbf{2}, \mathcal{F}]$ cannot be regarded as a tensor $\mathbf{2} \otimes \mathcal{F}$ in the 2-category $\mathfrak{Top}_{sk}$), but if we are given any 2-cell $\alpha \colon f \to g$ between morphisms $f, g \colon \mathcal{F} \rightrightarrows \mathcal{E}$ such that g is skeletal (in particular, any 2-cell of $\mathfrak{Top}_{sk}$ with source $\mathcal{F}$), the induced morphism $[\mathbf{2}, \mathcal{F}] \to \mathcal{E}$ is skeletal, since its composite with the open inclusion is g. Hence we may define γ on such 2-cells by lifting the morphism $[\mathbf{2}, \mathcal{F}] \to \mathcal{E}$ to a morphism $[\mathbf{2}, \gamma\mathcal{F}] \to \gamma\mathcal{E}$. Thus we have shown that γ is a 2-functor $\mathfrak{Top}_{sk} \to \mathfrak{Top}_{sk}$; and since the canonical map $\gamma\mathcal{E} \to \mathcal{E}$ is a natural transformation from γ to the identity, which is an equivalence precisely when $\mathcal{E}$ satisfies De Morgan's law, γ is right adjoint to the inclusion $\mathfrak{DMTop}_{sk} \to \mathfrak{Top}_{sk}$. □

In passing, we note that it is almost immediate from the definition of skeletal morphisms that the assignment $\mathcal{E} \mapsto \mathbf{sh}_{\neg\neg}(\mathcal{E})$ defines a right adjoint to the inclusion $\mathfrak{BooTop} \to \mathfrak{Top}_{sk}$, where $\mathfrak{BooTop}$ is the 2-category of Boolean toposes and geometric morphisms. (Of course, any geometric morphism between Boolean

toposes is skeletal.) Note also that $\gamma\mathcal{E}$ is Boolean only if $\mathcal{E}$ itself is Boolean, since in this case it must coincide with $\mathbf{sh}_{\neg\neg}(\mathcal{E})$ and hence the latter maps surjectively to $\mathcal{E}$.

Remark 4.6.13 An alternative proof of the 2-functoriality of γ may be given along the following lines. Given a skeletal morphism $f\colon \mathcal{F} \to \mathcal{E}$, there is a unique morphism $\theta\colon f^*((\Omega_{\neg\neg})_{\mathcal{E}}) \to (\Omega_{\neg\neg})_{\mathcal{F}}$ making the diagram

$$\begin{tikzcd}
f^*(2_{\mathcal{E}}) \arrow[r, "\sim"] \arrow[d, tail] & 2_{\mathcal{F}} \arrow[d, tail] \\
f^*((\Omega_{\neg\neg})_{\mathcal{E}}) \arrow[r, "\theta"] & (\Omega_{\neg\neg})_{\mathcal{F}}
\end{tikzcd}$$

commute, since $f^*(2) \rightarrowtail f^*(\Omega_{\neg\neg})$ is $\neg\neg$-dense and $\Omega_{\neg\neg}$ is a $\neg\neg$-sheaf. (Note that this morphism is *not* simply the restriction of the familiar comparison map $\phi_1\colon f^*(\Omega_{\mathcal{E}}) \to \Omega_{\mathcal{F}}$; the latter does not map $f^*(\Omega_{\neg\neg})$ into $\Omega_{\neg\neg}$ in general.) θ is easily seen to be a Boolean homomorphism (since it is homomorphic on the $\neg\neg$-dense subalgebra $f^*(2_{\mathcal{E}})$), and hence a morphism of sites when its domain and codomain are equipped with their coherent coverages; hence it corresponds to a geometric morphism $\gamma\mathcal{F} \to f^*(\gamma\mathcal{E})$ over $\mathcal{F}$, by C2.3.4 and C2.4.3. Composing this with the pullback of f along $\gamma\mathcal{E} \to \mathcal{E}$ yields the effect of γ on a 1-cell of $\mathfrak{Top}_{sk}$; to define its effect on 2-cells, we use the fact that $\gamma\mathcal{E} \to \mathcal{E}$ is a fibration in $\mathfrak{Top}$, by B4.4.11. Thus, given a 2-cell $\alpha\colon f \to g$ with codomain $\mathcal{E}$, we obtain a diagram

$$\begin{tikzcd}
f^*(\gamma\mathcal{E}) \arrow[r] & \gamma\mathcal{E} \\
g^*(\gamma\mathcal{E}) \arrow[u] \arrow[ur] \arrow[u, phantom, shift right=8, "\downarrow"] &
\end{tikzcd}$$

commuting up to a (non-invertible) 2-cell, where the left vertical arrow is induced by the Boolean homomorphism $\alpha_{\Omega_{\neg\neg}}\colon f^*(\Omega_{\neg\neg}) \to g^*(\Omega_{\neg\neg})$ in $\mathcal{F}$. And if f and g are skeletal, then the diagram

$$\begin{tikzcd}
\gamma\mathcal{F} \arrow[r] \arrow[dr] & f^*(\gamma\mathcal{E}) \\
& g^*(\gamma\mathcal{E}) \arrow[u]
\end{tikzcd}$$

commutes in $\mathfrak{Top}/\mathcal{F}$, since the corresponding diagram of Boolean algebras in $\mathcal{F}$ commutes (by uniqueness of θ). So on pasting the two diagrams above, we obtain the effect of γ on the 2-cell α.

A natural 'companion' to 4.6.8, which was also first investigated by Gleason in the spatial case, is the projectivity theorem for extremally disconnected spaces; this was developed for toposes in [515]. First, we need an algebraic equivalent of De Morgan's law: given an internal lattice (or ring) A in a topos $\mathcal{E}$, we say an ideal J of A is *proper* (no connection with the notion of proper map!) if it satisfies $\neg(1_A \in J)$; by a *maximal ideal* we of course mean a maximal proper ideal. An ideal J of a lattice A is said to be *prime* if it is proper and satisfies

$$(\forall a_1, a_2 \colon A)((a_1 \wedge a_2 \in J) \Rightarrow ((a_1 \in J) \vee (a_2 \in J))) ;$$

primeness for ring ideals is defined similarly, with multiplication in place of meet.

Lemma 4.6.14 *For a topos $\mathcal{E}$, the following are equivalent:*

(i) *$\mathcal{E}$ satisfies De Morgan's law.*

(ii) *For any internal distributive lattice A in $\mathcal{E}$, every maximal ideal of A is prime.*

(iii) *For any internal commutative ring A in $\mathcal{E}$, every maximal ideal of A is prime.*

(iv) *For any internal Boolean algebra B in $\mathcal{E}$, the maximal ideals of B are exactly the prime ideals.*

Proof (i) $\Rightarrow$ (ii): Let A be an internal distributive lattice in $\mathcal{E}$, J a maximal ideal of A, and suppose given elements a_1, a_2 satisfying $(a_1 \wedge a_2 \in J)$. Let K_i $(i = 1, 2)$ be the ideal generated by J together with a_i, i.e. the ideal of all elements of the form $b \vee c$ where $b \in J$ and $c \leq a_i$. By distributivity, it is clear that $K_1 \cap K_2 = J$; so since J is proper we have $\neg((1_A \in K_1) \wedge (1_A \in K_2))$. Since De Morgan's law holds, this implies $\neg(1_A \in K_1) \vee \neg(1_A \in K_2)$, i.e. one of K_1 or K_2 must be proper. But since $J \subseteq K_i$ for $i = 1, 2$, this implies that we have $(J = K_1) \vee (J = K_2)$, and hence $(a_1 \in J) \vee (a_2 \in J)$. So we have shown that J is prime.

The proof of (i) $\Rightarrow$ (iii) is identical apart from notational changes. For (ii) $\Rightarrow$ (iv) and (iii) $\Rightarrow$ (iv), we note that a Boolean algebra is both a distributive lattice and a commutative ring (and the two senses of the word 'ideal' coincide), so either (ii) or (iii) implies that maximal ideals in a Boolean algebra are prime. But the converse holds without any assumption on $\mathcal{E}$: for if J is a prime ideal in a Boolean algebra B, then for any $b \in B$ we have $(b \in J) \vee (\neg b \in J)$, and so if K is any proper ideal containing B, then we have

$$\begin{aligned} (b \in K) &\Rightarrow \neg(\neg b \in K) \text{ since } K \text{ is proper} \\ &\Rightarrow \neg(\neg b \in J) \text{ since } J \subseteq K \\ &\Rightarrow (b \in J) \end{aligned}$$

so $K = J$.

(iv) $\Rightarrow$ (i): Consider the internal Boolean algebra $\Omega_{\neg\neg}$. This has a unique proper ideal $\{\bot\}$, as we saw in the proof of 4.6.8, so the latter must be maximal.

But the assertion that $\bot$ is a prime ideal of $\Omega_{\neg\neg}$ is exactly the statement of De Morgan's law (restricted to $\neg\neg$-stable truth-values, but this is sufficient since $\neg(p \wedge q)$ is equivalent to $\neg(\neg\neg p \wedge \neg\neg q)$). □

To deduce the projectivity theorem for toposes satisfying De Morgan's law, we shall have to make use of Zorn's Lemma in the form in which we stated it in 4.5.14. We must therefore restrict ourselves to toposes which are localic over the 'classical' topos **Set**, in which the axiom of choice is presumed to hold.

Theorem 4.6.15 *For a localic* **Set***-topos* $\mathcal{E}$*, the following are equivalent:*

(i) *$\mathcal{E}$ satisfies De Morgan's law.*

(ii) *$\mathcal{E}$ is projective with respect to entire surjections in* $\mathfrak{Top}/\mathbf{Set}$.

(iii) *$\mathcal{E}$ is projective with respect to proper separated localic surjections in* $\mathfrak{Top}/\mathbf{Set}$.

(iv) *Every compact Hausdorff locale in $\mathcal{E}$ which is consistent (in the sense of* C2.4.8*) has a point.*

(v) *Every coherent internal locale in $\mathcal{E}$ which is consistent has a point.*

Proof (i)$\Rightarrow$(v): Let X be a coherent internal locale in $\mathcal{E}$: that is, $\mathcal{O}(X) \cong I(B)$, where B is the distributive lattice of compact open sublocales of X. It is clear that such an X is consistent iff B is non-degenerate, i.e. satisfies $\neg(0_B = 1_B)$. Also, points of X correspond to completely prime filters in $\mathcal{O}(X)$ (cf. C1.2.2); and since every element of $I(B)$ is a join of principal ideals, it is easy to see that a completely prime filter of $I(B)$ is determined by the principal ideals which it contains, and that in fact there is a bijection between completely prime filters in $I(B)$ and (ordinarily) prime filters in B. But if B is non-degenerate then the poset of proper filters in B is (inhabited, and hence) inductive; so by 4.5.14 it contains a maximal element. And by 4.6.14 (applied to the dual lattice B^{op}) such a maximal filter is prime.

(v)$\Rightarrow$(iv) follows easily from C4.1.13, where we showed that every compact Hausdorff locale is a retract of a coherent locale: clearly, the latter is necessarily consistent if the former is.

(iv)$\Rightarrow$(iii): Since proper separated localic surjections are stable under pullback in $\mathfrak{Top}$, it suffices to prove that every such map with codomain $\mathcal{E}$ can be split. But such a map with codomain $\mathcal{E}$ corresponds to a consistent compact Hausdorff locale in $\mathcal{E}$, and a splitting of it corresponds to a point of this locale; so the implication is immediate. And (iii)$\Rightarrow$(ii) is immediate since entire morphisms are localic, proper and separated.

(ii)$\Rightarrow$(i): Suppose $\mathcal{E}$ is projective. Then the Gleason cover map $\gamma\mathcal{E} \to \mathcal{E}$ must split. But this is equivalent to saying that the internal Boolean algebra $\Omega_{\neg\neg}$ in $\mathcal{E}$ has a prime ideal, and we saw in the proof of 4.6.14 that this is equivalent to De Morgan's law in $\mathcal{E}$. (Alternatively, we can use 4.6.7(ii) to deduce that $\mathcal{E}$ inherits De Morgan's law from $\gamma\mathcal{E}$.) □

Remark 4.6.16 In combination with 4.6.8, the result of 4.6.15 implies that any geometric morphism $f: \mathcal{F} \to \mathcal{E}$ between localic **Set**-toposes can be lifted to a morphism $\gamma\mathcal{F} \to \gamma\mathcal{E}$. However, in view of the arbitrary choice involved in the lifting, we cannot hope to make the latter functorial on a larger category than that of skeletal maps.

Suggestions for further reading: Johnstone [509, 510, 512, 515], Kock & Reyes [645], Mulvey & Pelletier [882], Niefield & Rosenthal [897, 898].

D4.7 Real numbers in a topos

The familiar set-theoretic constructions of the ring of integers, the field of rationals and the field of real numbers starting from the semiring of natural numbers may all be carried out within the higher-order type theory described in Section D4.3, and so make sense in any topos $\mathcal{E}$, provided it has a natural number object from which the construction can start. However, the non-classical nature of the internal logic of $\mathcal{E}$ makes itself felt at the final stage of this construction: although all sensible constructions of the integers and of the rationals yield the same results, there are various different constructions of the reals which are classically equivalent but may yield different results in a non-Boolean topos.

Throughout this section we work in a topos with a natural number object N. (The constructions from N of the object Z of integers and the object Q of rationals, being essentially first-order, could in fact be carried out in any pretopos with a natural number object.) We define Z to be the quotient of $N \times N$ by the equivalence relation which identifies $\langle m, n\rangle$ with $\langle m', n'\rangle$ iff $m + n' = m' + n$, and we define a monomorphism $N \rightarrowtail Z$ to send n to the equivalence class of $\langle n, o\rangle$. The arithmetic operations of addition and multiplication (cf. A2.5.4) are extended to Z in an obvious way, and make it into a commutative ring: in fact it is the free commutative ring (with 1) on no generators, and its underlying additive group is the free group on one generator in $\mathcal{E}$ (cf. 5.3.5 below). Then we define Q to be the field of fractions of Z, i.e. the quotient of $Z \times Z*$ (where $Z*$ is the subobject of nonzero elements of Z, i.e. the complement of $0: 1 \rightarrowtail Z$ – note that this element is indeeed complemented, since the equivalence relation we defined on $N \times N$ is complemented and hence Z is decidable) by the equivalence relation which identifies $\langle z, w\rangle$ with $\langle z', w'\rangle$ iff $z \cdot w' = z' \cdot w$. Once again, this equivalence relation is complemented, and so Q is decidable; we embed Z into Q by sending z to the equivalence class of $\langle z, 1\rangle$. And we can again extend the arithmetic operations from Z to Q: for example, the sum of the equivalence classes of $\langle z, w\rangle$ and $\langle z', w'\rangle$ is the equivalence class of $\langle zw' + z'w, ww'\rangle$. The total ordering on N, which we defined in A2.5.13, also extends to both Z and Q in a straightforward manner.

Lemma 4.7.1 *In any topos with a natural number object, Q with its arithmetic operations is a model of the coherent theory of fields described in* 1.1.7(h); *and*

Q with its order relation is a model of the theory $\mathbb{L}_\infty$ of dense linearly ordered objects without endpoints, described in 3.4.11.

Proof It is straightforward to verify that the usual set-theoretic proofs of these facts are constructive, and so valid in any topos. □

We note that the constructions used so far are all preserved by coherent functors, in particular by inverse image functors: since the latter also preserve natural number objects by A2.5.6, it follows that they preserve objects of integers and of rationals. In particular, in any topos of the form $[\mathcal{C}, \mathbf{Set}]$ the object of rationals is simply the constant functor on $\mathcal{C}$ with value $\mathbb{Q}$, and in a topos of the form $\mathbf{Sh}(X)$ it is the constant sheaf with stalk $\mathbb{Q}$, i.e. the sheaf whose sections over an open U are locally constant functions $U \to \mathbb{Q}$ (equivalently, continuous functions $U \to \mathbb{Q}$, where $\mathbb{Q}$ is given the discrete topology).

For most of this section, we shall consider real numbers as Dedekind sections of the object of rationals (though we briefly discuss an alternative construction in 4.7.12 below). However, there are still various choices to be made: do we work with 'one-sided' or 'two-sided' sections, and if we use two-sided sections how do we express the idea that the two sides are at zero distance from each other? Classically, these choices do not matter, since they yield isomorphic models of the real numbers; but we shall see that in constructive logic they yield different results.

We begin with one-sided sections. Let $(A, <)$ be a model of the theory $\mathbb{L}_\infty$ in a topos $\mathcal{E}$: we define an *open lower section* of A to be a subobject $L \rightarrowtail A$ such that the sequents

$$(((x < y) \wedge (y \in \ulcorner L \urcorner)) \vdash_{x,y} (x \in \ulcorner L \urcorner)),$$

$$((x \in \ulcorner L \urcorner) \vdash_x (\exists y)((x < y) \wedge (y \in \ulcorner L \urcorner))) \text{ and}$$

$$(\top \vdash (\exists x)(x \in \ulcorner L \urcorner))$$

are satisfied. We say an open lower section is *proper* if it also satisfies

$$(\top \vdash (\exists x)\neg(x \in \ulcorner L \urcorner)) :$$

the reason for separating this axiom from the other three is that, unlike them, it is not expressible as a coherent sequent. Clearly, we may form a subobject $A_\downarrow^+$ of PA whose elements are the open lower sections of A; and we shall write $A_\downarrow$ for the subobject of $A_\downarrow^+$ whose elements are the proper sections. Associated with any element $a \colon 1 \to A$, we have a proper open lower section $\downarrow a = \{x \colon A \mid x < a\}$; this defines a morphism $\downarrow\colon A \to A_\downarrow$. We define the object R_l of *lower semicontinuous real numbers* in $\mathcal{E}$ to be $Q_\downarrow$.

We may also consider open lower sections of A as models of a propositional geometric theory over $\mathcal{E}$ (in the sense of B4.2.12, or of 1.1.7(m)), expressed in a language with primitive propositions of the form $(a \in \ulcorner L \urcorner)$, $a \in A$. For simplicity we shall assume that $\mathcal{E}$ is defined over $\mathbf{Set}$, and that A is of the form

$\gamma^* A_0$ for some $\mathbb{L}_\infty$-model A_0 in **Set** (as it is in the case of interest to us), where $\gamma \colon \mathcal{E} \to \mathbf{Set}$ is the geometric morphism: then we may identify open lower sections of A in $\mathcal{E}$ with geometric morphisms from $\mathcal{E}$ to the classifying topos of this propositional theory.

Lemma 4.7.2 *Let A be an $\mathbb{L}_\infty$-model in* **Set**, *and let $\mathbb{T}$ denote the propositional theory of open lower sections of A. Then the classifying topos $\mathbf{Set}[\mathbb{T}]$ may be taken to be $\mathbf{Sh}(Y)$, where Y is the locale whose frame $\mathcal{O}(Y)$ is generated by symbols $\uparrow a$, $a \in A$, subject to the relations $\uparrow a \wedge \uparrow b = \uparrow(\max\{a, b\})$,*

$$\uparrow a = \bigvee \{\uparrow b \mid b > a\}$$

for all $a \in A$, and $\top = \bigvee\{\uparrow a \mid a \in A\}$. Moreover, this locale is spatial; it may be identified with the space $Y = \overline{A}^+$ obtained from the (classical) Dedekind–MacNeille completion $\overline{A}$ of A by adding a top point ∞, endowed with the topology generated by the sets $\{y \in Y \mid a < y\}$, $a \in A$.

Proof By 3.1.14 (or by B4.2.12), we know that $\mathbf{Set}[\mathbb{T}]$ is localic over **Set**, and the frame corresponding to it is the Lindenbaum algebra of $\mathbb{T}$. The latter is generated by the primitive propositions $(a \in \ulcorner L \urcorner)$, $a \in A$, subject to relations corresponding to the axioms of the theory; the relations in the statement of the lemma are a direct translation of the three coherent sequents we wrote down earlier, with a change in notation from $(a \in \ulcorner L \urcorner)$ to $\uparrow a$ throughout. So the first assertion of the lemma is a triviality. The second assertion, that the locale Y is spatial, is slightly less trivial; but points of Y correspond to frame homomorphisms $\mathcal{O}(Y) \to \mathcal{O}(1)$, and hence to mappings $\{\uparrow a \mid a \in A\} \to \mathcal{O}(1)$ preserving the given relations; and it is easy to see that these correspond exactly to the points of $\overline{A}$ (the point y which corresponds to a lower section L sending $\uparrow a$ to $\top$ iff $a \in L$), together with the point ∞ which sends every $\uparrow a$ to $\top$. Moreover, the topology on Y_p is that generated by the sets $\{y \in Y \mid a < y\}$; we must show that this is isomorphic to $\mathcal{O}(Y)$.

However, the first relation in the definition of $\mathcal{O}(Y)$ (plus the infinite distributive law) tells us that any element of $\mathcal{O}(Y)$ is equal to a join of generators; and the other relations tell us that we have $\uparrow a \leq \bigvee\{\uparrow b \mid b \in S\}$ (resp. $\top \leq \bigvee\{\uparrow b \mid b \in S\}$) provided that for every $c > a$ (resp. for every $c \in A$) there exists $b \in S$ with $b < c$. And these relations are in turn implied by the corresponding inclusions between subsets of Y_p; so the canonical frame surjection $\mathcal{O}(Y) \twoheadrightarrow \mathcal{O}(Y_p)$ is bijective. □

Note that if we omit the top point ∞ from the space Y of 4.7.2, we obtain one which is not sober. This is related to the fact that the 'properness' axiom for open lower sections cannot be formulated geometrically.

Corollary 4.7.3 *Let X be a (sober) topological space. Then, in the topos* $\mathbf{Sh}(X)$,

(i) *the object R_l^+ is the sheaf whose sections over an open $V \subseteq X$ are all lower semicontinuous functions from V to the set $\mathbb{R}^+ = \mathbb{R} \cup \{\infty\}$;*

(ii) R_l *is the sheaf whose sections over an open* V *are all those lower semicontinuous functions* $f: V \to \mathbb{R}$ *which are 'locally bounded above', i.e. such that for all* $x \in V$ *there exists* $q \in \mathbb{Q}$ *and a neighbourhood* W *of* x *such that* $f(y) < q$ *for all* $y \in W$.

Proof (i) By definition, sections of R_l^+ over V correspond to models of the geometric propositional theory of open lower sections of $\mathbb{Q}$ in the topos $\mathbf{Sh}(V)$. So this is immediate from the fact that lower semicontinuous functions are exactly the continuous functions when $\mathbb{R}^+$ is equipped with the topology described in 4.7.2, plus the fact that continuous functions $V \to \mathbb{R}^+$ correspond bijectively to (isomorphism classes of) geometric morphisms $\mathbf{Sh}(V) \to \mathbf{Sh}(\mathbb{R}^+)$ (C1.4.5).

(ii) follows from (i) and the fact that, for the lower section L corresponding to a lower semicontinuous $f: V \to \mathbb{R}^+$, the truth-value of $(\exists x)\neg(x \in \ulcorner L \urcorner)$ is the union over all $q \in \mathbb{Q}$ of the interiors of the sets $\{x \in V \mid f(x) < q\}$. □

The local boundedness condition in (ii) is stronger than simply asserting that f never takes the value ∞; for example, consider the function f defined on $\mathbb{R}$ by $f(x) = 0$ for all $x \leq 0$ and $f(x) = 1/x$ for $x > 0$. For more information about the approach to real numbers via one-sided Dedekind sections, see [1001].

One obvious problem with one-sided sections is that they fail to inherit the full algebraic structure of $\mathbb{Q}$: although we can add two lower semicontinuous functions, we do not have additive inverses, and multiplication is even more problematic (unless we restrict ourselves to functions taking only positive values). Therefore, for most applications, it is more convenient to work with two-sided sections; that is, with pairs $\langle L, U \rangle$ such that L is an open lower section of our $\mathbb{L}_\infty$-model A, U is an open upper section, L and U are disjoint (i.e.

$$(((x \in \ulcorner L \urcorner) \wedge (x \in \ulcorner U \urcorner)) \vdash_x \bot) \ )$$

and 'L and U are at zero distance apart'. However, we still have a choice about how we formulate the last of these conditions. If we wish our sections to be models of a geometric theory, we are led to use the axiom

$$((x < y) \vdash_{x,y} ((x \in \ulcorner L \urcorner) \vee (y \in \ulcorner U \urcorner))) \ .$$

However, we can also consider the pair of axioms

$$(((x < y) \wedge \neg(y \in \ulcorner U \urcorner)) \vdash_{x,y} (x \in \ulcorner L \urcorner)) \text{ and}$$

$$(((x < y) \wedge \neg(x \in \ulcorner L \urcorner)) \vdash_{x,y} (y \in \ulcorner U \urcorner)) \ ;$$

these are both strictly weaker than the coherent sequent above, and (in the presence of the other axioms) they express the idea that L is the largest open lower section disjoint from U, and U is the largest open upper section disjoint from L. We shall call a pair $\langle L, U \rangle$ a *Dedekind section* of A if it satisfies the stronger axiom above (plus the earlier axioms), and a *MacNeille section* if it

satisfies the two weaker ones. Clearly, we have subobjects $\overline{A}_d$ and $\overline{A}_m$ of $PA \times PA$ corresponding to these two notions; in the particular case when $A = Q$, we shall write R_d and R_m. Note also that we always have $A_d \subseteq A_m$ as subobjects of $PA \times PA$; and we have an embedding $A \rightarrowtail A_d$ given by $x \mapsto \langle \downarrow x, \uparrow x \rangle$, as well as an embedding $A_m \rightarrowtail A_\downarrow$ given by $\langle L, U \rangle \mapsto L$. (The latter map is monic because, for a MacNeille section $\langle L, U \rangle$, we can recover U from L:

$$((y \in \ulcorner U \urcorner) \dashv\vdash_y (\exists x)((x < y) \wedge \neg(x \in \ulcorner L \urcorner))) \, .)$$

As with the one-sided sections, we begin our investigation of two-sided sections by considering the classifying topos for the geometric propositional theory of Dedekind sections of an $\mathbb{L}_\infty$-model A in **Set**. For simplicity, we shall restrict ourselves to the case $A = \mathbb{Q}$ (recall that this is not as special as it seems, since every countable model of $\mathbb{L}_\infty$ is isomorphic to $\mathbb{Q}$).

Lemma 4.7.4 *Let $\mathbb{T}_d$ denote the theory of Dedekind sections of $\mathbb{Q}$. The classifying topos* $\mathbf{Set}[\mathbb{T}_d]$ *is of the form* $\mathbf{Sh}(Y)$, *where Y is the locale corresponding to the frame $\mathcal{F}(\mathbb{R})$ (sometimes called the frame of* formal reals*) which is generated by symbols (p, q) where p and q are rationals with $p < q$, subject to the following relations:*

$$(p_1, q_1) \wedge (p_2, q_2) = \begin{cases} (\max\{p_1, p_2\}, \min\{q_1, q_2\}) & \textit{if } \max\{p_1, p_2\} < \min\{q_1, q_2\} \\ \bot & \textit{otherwise;} \end{cases}$$

$$(p, q) = (p, q_1) \vee (p_1, q) \quad \textit{if } p \leq p_1 < q_1 \leq q;$$

$$(p, q) = \bigvee\{(p', q') \mid p < p' < q' < q\} \quad ; \textit{ and}$$

$$\top = \bigvee\{(p, q) \mid p < q\} \, .$$

Moreover, this locale is spatial, and it may be identified with the space $\mathbb{R}$ equipped with the usual (Euclidean) topology.

Proof As in 4.7.2, the first part is merely an explicit description of the Lindenbaum algebra of $\mathbb{T}_d$. The latter is generated by primitive propositions of the forms $(p \in \ulcorner L \urcorner)$ and $(q \in \ulcorner U \urcorner)$ $(p, q \in \mathbb{Q})$; the generators (p, q) which we have used for $\mathcal{F}(\mathbb{R})$ are simply the conjunctions $((p \in \ulcorner L \urcorner) \wedge (q \in \ulcorner U \urcorner))$. (For some purposes it is convenient to allow the use of the symbol (p, q) when $p \geq q$; of course, in this case, it is equal to the bottom element of the frame, by the disjointness axiom.) We can recover the original generator $(p \in \ulcorner L \urcorner)$ as the join $\bigvee\{(p, q) \mid q > p\}$, and similarly for $(q \in \ulcorner U \urcorner)$; so the elements (p, q) do indeed generate the frame. Then the first condition in our presentation becomes a translation of the axioms saying that L is a lower section and U is an upper section; the third translates the condition that they are open, and the fourth is the condition that they are both inhabited. Finally, the second condition expresses the 'Dedekind version' of the condition that they are at zero distance apart.

The real meat of the lemma, once again, comes in the assertion that the locale defined by this frame presentation is spatial. Of course, the points of the locale may be identified with $\mathbb{T}_d$-models in **Set**, that is with real numbers (Dedekind sections of $\mathbb{Q}$) in the classical sense; and the topology induced upon them is that generated by the rational open intervals $\{x \in \mathbb{R} \mid p < x < q\}$, which are the images of the formal symbols (p,q) – that is, the usual topology on $\mathbb{R}$. The problem is thus to show that the subframe of $P(\mathbb{R})$ generated by these subsets satisfies only those conditions which are deducible from those in the given presentation of $\mathcal{F}(\mathbb{R})$. We note that the first condition in the presentation enables us to reduce any element of $\mathcal{F}(\mathbb{R})$ to a (possibly empty) join of generators, so as in 4.7.2 we are reduced to considering inequalities of the form $(p,q) \leq \bigvee\{(p_i,q_i) \mid i \in I\}$ or $\top \leq \bigvee\{(p_i,q_i) \mid i \in I\}$. We shall deal only with the first of these; the second is similar.

The key ingredient that we need is the Heine–Borel theorem; that is, the assertion that every bounded closed interval in $\mathbb{R}$ is compact. Given this, we see that if an actual interval $\{x \in \mathbb{R} \mid p < x < q\}$ is covered by the intervals $\{x \mid p_i < x < q_i\}$, then for any (p',q') with $p < p' < q' < q$ we must have a finite subcover of the closed interval $\{x \mid p' \leq x \leq q'\}$, and hence we can find a finite sequence of indices $i_1, \ldots, i_n$ with $p_{i_1} < p'$, $q_{i_n} > q'$ and $p_{i_{j+1}} < q_{i_j}$ for all j. But these inequalities, plus the second and third conditions in the definition of $\mathcal{F}(\mathbb{R})$, ensure that the inequality $(p,q) \leq \bigvee\{(p_i,q_i) \mid i \in I\}$ holds in the latter. □

The use of the Heine–Borel theorem in the foregoing proof is unavoidable. Although this theorem does not make any use of the axiom of choice, it does require the law of excluded middle; we shall see in 4.7.13 below that there are toposes in which it fails for the internal space of Dedekind reals. On the other hand, the locale R_f of formal reals (the locale corresponding to the frame $\mathcal{F}(\mathbb{R})$, as constructed in a topos) is constructively locally compact; more specifically, for any two rationals p and q, the quotient frame $\mathcal{F}([p,q])$ of $\mathcal{F}(\mathbb{R})$ obtained by adjoining the relations $(r,s) = \bot$ whenever $s \leq p$ or $r \geq q$ (the 'formal closed interval $[p,q]_f$') is compact (see [520, IV 1.2], for a proof). Hence, if we construct the locale R_f inside the topos of 4.7.13, it will not be spatial.

Corollary 4.7.5 *Let X be a (sober) topological space. Then, in the topos* $\mathbf{Sh}(X)$,

(i) *R_d is the sheaf whose sections over an open $V \subseteq X$ are all continuous functions $f \colon V \to \mathbb{R}$;*

(ii) *R_m is the sheaf whose sections over V are pairs $\langle \underline{f}, \overline{f} \rangle$ of functions $V \to \mathbb{R}$ satisfying*

$$\overline{f}(x) = \inf\{\sup\{\underline{f}(y) \mid y \in U\} \mid U \in \mathcal{N}_x\}$$

and

$$\underline{f}(x) = \sup\{\inf\{\overline{f}(y) \mid y \in U\} \mid U \in \mathcal{N}_x\}$$

for all $x \in V$, where $\mathcal{N}_x$ denotes the filter of neighbourhoods of x in V.

Proof (i) This follows from 4.7.4 just as 4.7.3(i) followed from 4.7.2.

(ii) A section of R_m over V is, in particular, a pair of proper one-sided sections of $\mathbb{Q}$ in $\mathbf{Sh}(V)$, and hence corresponds to a pair of functions $\underline{f}: V \to \mathbb{R}$ and $\overline{f}: V \to \mathbb{R}$ which are respectively lower semicontinuous and upper semicontinuous. The 'MacNeille form' of the condition that L and U are at zero distance apart translates easily into the conditions on $\underline{f}$ and $\overline{f}$ given in the statement; but these conditions imply the required semicontinuity properties of the two functions, and so we do not have to state them explicitly. □

Of course, a pair of functions $(\underline{f}, \overline{f})$ as in 4.7.5(ii) corresponds to an element of $R_d(V) \subseteq R_m(V)$ iff the two functions coincide (in which case they are necessarily two-sidedly continuous). On the other hand, it is easy to give examples of MacNeille reals which are not Dedekind reals; for example, the pair of functions $(\underline{f}, \overline{f}): \mathbb{R} \rightrightarrows \mathbb{R}$ defined by $\underline{f}(x) = \overline{f}(x) = 0$ for all $x < 0$, $\underline{f}(x) = \overline{f}(x) = 1$ for $x > 0$, $\underline{f}(0) = 0$ and $\overline{f}(0) = 1$. We shall revert to the question of when R_d and R_m coincide in 4.7.11 below.

Our next task, however, is to extend the order and algebraic structures from Q to R. This may be done either for R_d or for R_m, by essentially the same constructions; although we are mainly interested in R_d, we shall give the constructions for R_m, since they may then easily be restricted to R_d. If t is a term of type R_m or R_d, we shall write L_t and U_t for the terms of type PQ which should properly be written $\mathsf{fst}(\ulcorner i \urcorner(t))$ and $\mathsf{snd}(\ulcorner i \urcorner(t))$, where $i: R \rightarrowtail PQ \times PQ$ is the canonical inclusion. We deal first with the order structure: we have two ways of defining an order relation on MacNeille sections, the *strong ordering*

$$((x < y) \dashv\vdash (\exists p: Q)((p \in U_x) \wedge (p \in L_y)))$$

and the *weak ordering*

$$((x \leq y) \dashv\vdash (\forall p: Q)((p \in L_x) \Rightarrow (p \in L_y))) .$$

(The apparent asymmetry in the definition of $\leq$ is only apparent: because U_x is recoverable from L_x, as we noted earlier, the condition $L_x \subseteq L_y$ is equivalent to $U_y \subseteq U_x$.) It is straightforward to verify that both these relations are transitive, that the first is irreflexive, and that the second is reflexive and antisymmetric. Also, the first relation is a two-sided ideal in the second (that is, $(x < y)$ implies $(x \leq y)$, and $(x \leq y < z \leq w)$ implies $(x < w)$); once again, this is easily verified from the definitions. And for variables x, y, z of type R_d (though not, in general, of type R_m) we have

$$((x < y) \vdash_{x,y,z} ((x < z) \vee (z < y))) ;$$

for from $(x < y)$ and the openness of the sections defining x and y we can deduce the existence of rationals $p < q$ with $((p \in U_x) \wedge (q \in L_y))$, and then we must have either $(p \in L_z)$ or $(q \in U_z)$.

On the other hand, it is not true that $(x \leq y)$ is equivalent to $((x < y) \vee (x = y))$, nor do we have $(\forall x, y)((x \leq y) \vee (y \leq x))$, even for Dedekind reals x and y. For a counterexample to the first in $\mathbf{Sh}(\mathbb{R})$, take x to be the identity function $\mathbb{R} \to \mathbb{R}$ and y to be the absolute-value function; then the truth-value of $(x \leq y)$ is the whole of $\mathbb{R}$, but that of $((x < y) \vee (x = y))$ is $\mathbb{R} \setminus \{0\}$. Similarly, for the second we take x to be the identity function and y to be the function $t \mapsto -t$ (or the constant function with value 0). However, we do have one useful relationship between the two orderings:

Lemma 4.7.6 *In any topos, we have $((x \leq y) \dashv\vdash_{x,y} \neg(y < x))$ for variables x, y of type R_m. Hence in particular the objects R_m and R_d are $\neg\neg$-separated in any topos.*

Proof Let x and y be MacNeille reals. Clearly, from $(x \leq y)$ and $(y < x)$ we deduce $(\exists p)((p \in U_y) \wedge (p \in L_y))$, contradicting the disjointness axiom; so $(x \leq y)$ implies $\neg(y < x)$. Conversely, $\neg(y < x)$ is equivalent to $(\forall p)\neg((p \in U_y) \wedge (p \in L_x))$, so from $(q \in L_x)$ we deduce $(\exists p)((p > q) \wedge (p \in L_x))$ by openness, and hence $(\exists p)((p > q) \wedge \neg(p \in U_y))$. But the latter implies $(q \in L_y)$ by the definition of a MacNeille real; so we have verified that $L_x \subseteq L_y$, i.e. that $(x \leq y)$.

The second assertion follows immediately from the first and A4.3.6(a), since the formula $(x = y)$ whose interpretation is the diagonal subobject of $R_m \times R_m$ (or of $R_d \times R_d$) is equivalent to $((x \leq y) \wedge (y \leq x))$ and hence also equivalent to $\neg((x < y) \vee (y < x))$. □

We shall sometimes write $(x \# y)$ for the relation $((x < y) \vee (y < x))$ on R_d, and call it the *apartness relation* on R_d. In general, a symmetric binary relation $\#$ on an object X is called an apartness relation if it satisfies

$$((x = y) \dashv\vdash_{x,y} \neg(x \# y))$$

and

$$((x \# y) \vdash_{x,y,z} ((x \# z) \vee (y \# z)))\ .$$

(For R_d, the second of these conditions follows from the condition $((x < y) \vdash_{x,y,z} ((x < z) \vee (z < y)))$, which we established earlier.) Objects with apartness relations have been extensively studied in the context of constructive mathematics, as a useful generalization of decidable objects (cf. [90, p. 22]); the fact that R_d comes equipped with a 'natural' apartness relation means that results established for such objects can be applied to it.

On the other hand, neither R_m nor R_d is decidable, in general: once again, a counterexample to $((x = y) \vee \neg(x = y))$ may be given in $\mathbf{Sh}(\mathbb{R})$ by taking x and y to be the identity and absolute value functions. Nor is either object a $\neg\neg$-sheaf, in general. For if we weaken the definition of a two-sided section of Q by replacing the positive condition that L and U should be inhabited with the negative condition that they should be nonempty (i.e. $\neg(\forall q \colon Q)\neg(q \in \ulcorner L \urcorner)$, etc.), we obtain objects $*R_d$ and $*R_m$ in which R_d and R_m are $\neg\neg$-densely embedded

(and which are still $\neg\neg$-separated, by the same proof as above), but which do not coincide with R_d and R_m in general; for example, in $\mathbf{Sh}(\mathbb{R})$ the function $(x \mapsto 1/|x|)$ (undefined at $x = 0$) yields a global section of $*R_d$ which does not lie in R_d or even in R_m.

In general, R_d enjoys more good properties than R_m: we have already seen that it has an apartness relation, and we shall see shortly that it has better algebraic properties as well. However, there is one respect in which R_m is better, namely order-completeness. Of course, neither R_d nor R_m is order-complete in the sense in which we usually use that term, since they lack top and bottom elements. One way to get round this would be to consider a closed interval $[a, b]$ instead of the whole real line; alternatively we can modify the definition of completeness. We shall say that a poset X is *conditionally complete* if every inhabited subset of X which has an upper bound has a least upper bound; this condition is also sometimes known as Dedekind-completeness, but since R_d does not satisfy it in a general topos that name would be confusing in the present context.

Lemma 4.7.7 *In any topos, R_m is a conditionally complete poset.*

Proof Let S be a bounded inhabited subset of R_m. To construct a least upper bound $x = \langle L, U\rangle$ for S, we take U to be the interior of the intersection of all the upper sections of the members of S, i.e.

$$U = \{q \colon Q \mid (\exists p < q)(\forall y \colon S)(p \in U_y)\} \; .$$

For L, we might be tempted to take simply the union of the lower sections of the members of S, but this will not yield a MacNeille real in general, so we have to define it from U:

$$L = \{p \colon Q \mid (\exists q \colon Q)((p < q) \wedge \neg(q \in U))\} \; .$$

Verification of most of the conditions for $\langle L, U\rangle$ to be a MacNeille real is straightforward (for example, inhabitedness of L and of U follow respectively from the conditions that S is inhabited and that it is bounded above – note that L certainly contains the union of the left sets of the members of S). The only one which requires some work is

$$(((p < q) \wedge \neg(p \in L)) \Rightarrow (q \in U)) \; ;$$

given p and q satisfying the hypotheses of this implication, let $r = (2p+q)/3$ and $s = (p + 2q)/3$, so that $p < r < s < q$. Then from the definition of L we obtain $\neg\neg(r \in U)$. But since $U \subseteq U_y$ for all $y \in S$, we have $U \cap L_y = \emptyset$ for all such y, and so this implies $\neg(r \in L_y)$ for all such y. Hence we have $(\forall y \colon S)(s \in U_y)$, and so $q \in U$, as required.

We have also observed that $L_y \subseteq L$ for all $y \in S$, so x is an upper bound for S. But if z is any upper bound for S, then its upper section U_z is open and contained in U_y for all $y \in S$, so it must be contained in U; hence $x \leq z$. □

It follows straightforwardly from 4.7.7 that any bounded closed interval $[a,b]_m \subseteq R_m$ is a complete poset in the usual sense. However, we shall see in 4.7.11 below that R_d is not conditionally complete in general. (On the other hand, we note for future reference that R_d does have binary join and meet operations – we shall denote them by $\max\{x,y\}$ and $\min\{x,y\}$, as is customary – which are obtained simply by taking unions and intersections of the appropriate sections.)

Next we consider the algebraic structure on R_d and R_m. In fact the same definitions work in both cases; until further notice, we shall cease to specify whether we are working with Dedekind or MacNeille reals, and write R to denote either R_d or R_m. Defining addition is straightforward: we set

$$\begin{aligned}&\langle L_1, U_1\rangle + \langle L_2, U_2\rangle \\ &\quad = \langle \{q_1 + q_2 \mid (q_1 \in L_1) \wedge (q_2 \in L_2)\}, \{q_1 + q_2 \mid (q_1 \in U_1) \wedge (q_2 \in U_2)\}\rangle\end{aligned}$$

and verify that this is a Dedekind (resp. MacNeille) real if both $\langle L_1, U_1\rangle$ and $\langle L_2, U_2\rangle$ are. It is also easy to verify that this operation makes R into an abelian group, the additive inverse of $\langle L, U\rangle$ being $\langle\{-q \mid q \in U\}, \{-q \mid q \in L\}\rangle$. However, multiplication is a little more tricky, because the classical subdivision into cases according to the signs of the numbers being multiplied will not work in our constructive context, where we do not have a 'trichotomy law' for the ordering. To get round this difficulty, we use an idea based on J. H. Conway's recursive definition of multiplication [248], which was first suggested in [504]. We first define the product of a real number and a rational (that is, we make R into a module over Q), for which we may legitimately subdivide into cases since the ordering on Q satisfies the trichotomy law:

$$q \cdot x = \begin{cases} \langle\{q \cdot r \mid r \in L_x\}, \{q \cdot r \mid r \in U_x\}\rangle & \text{if } q > 0 \\ \langle\{q \cdot r \mid r \in U_x\}, \{q \cdot r \mid r \in L_x\}\rangle & \text{if } q < 0 \\ 0 & \text{if } q = 0\,. \end{cases}$$

It is not hard to verify that this does indeed satisfy the identities for a Q-module. Now we exploit the idea that if $q_1 - x_1$ and $q_2 - x_2$ have the same sign then $(x_1 - q_1)(x_2 - q_2)$ should be positive; that is, we define $L_{x_1 \cdot x_2}$ to be

$$\begin{aligned}&\{q \colon Q \mid (\exists q_1, q_2)((((q_1 \in L_{x_1}) \wedge (q_2 \in L_{x_2})) \vee ((q_1 \in U_{x_1}) \wedge (q_2 \in U_{x_2})))\wedge \\ &\qquad (q + q_1 \cdot q_2 \in L_{q_2 \cdot x_1 + q_1 \cdot x_2}))\}\end{aligned}$$

with a similar definition for $U_{x_1 \cdot x_2}$ (in which the rationals q_1 and q_2 are taken from the sections on opposite sides of x_1 and x_2). A large amount of tedious but straightforward verification, which we omit, allows us to conclude

Proposition 4.7.8 *The above definitions of addition and multiplication make R into a commutative ring.* □

Next, it is natural to ask in what sense, if any, the real number objects are fields. The example of a spatial topos, in which the stalks of the sheaf of continuous real-valued functions are well known to be local rings but not fields in general, shows that we cannot hope for even R_d to be a field in the coherent sense defined in 1.1.7(h), although we can hope to show that it is a local ring in the sense of 1.1.7(g). In fact both R_d and R_m are fields in a constructively weaker sense: we say that a commutative ring is a *residue field* if it satisfies the nontriviality axiom $((0 = 1) \vdash \bot)$ and the non-coherent sequent

$$(\neg(\exists y)(xy = 1) \vdash_x (x = 0)) .$$

(The reason for the name 'residue field' is that if we attempt to form the residue field of a local ring – that is, if we factor out the ideal of non-invertible elements – what we obtain is exactly a ring satisfying this axiom.)

In order to prove that R_d and R_m are residue fields, we need to determine their invertible elements. The key step in this is the following:

Lemma 4.7.9 *We have*

$$(\forall x, y\colon R)((x \cdot y > 0) \Rightarrow (((x > 0) \wedge (y > 0)) \vee ((x < 0) \wedge (y < 0)))) .$$

Proof The hypothesis that $0 \in L_{x \cdot y}$ tells us that we have either $q \in L_x$ and $r \in L_y$ with $q \cdot r \in L_{q\,y + r \cdot x}$, or else $q \in U_x$ and $r \in U_y$ satisfying a similar condition. The analysis of the two cases is similar; each proceeds by considering the possible signs of q and r, and we shall do only the first case. Clearly, if $q \geq 0$ and $r \geq 0$ then we have $x > 0$ and $y > 0$, so we need only consider the sub-cases when one of q and r is strictly less than 0. If $q < 0 = r$, then the condition becomes $0 \in L_{q \cdot y}$, which implies $0 \in U_y$; but this contradicts $0 = r \in L_y$. If $q < 0 < r$, then we have $s \in U_y$ and $t \in L_x$ such that $q \cdot r = q \cdot s + r \cdot t$, or equivalently $q \cdot (r - s) = r \cdot t$; but $q \cdot (r - s) > 0$ since both q and $(r - s)$ are negative, and hence $t > 0$. But this forces $x > 0$, since $t \in L_x$, and we already have $y > 0$ since $r \in L_y$. Finally, we need to deal with the case when both q and r are negative. In this case, we obtain $s \in U_y$ and $t \in U_x$ such that $q \cdot r = q \cdot s + r \cdot t$, from which we get $r \cdot t = q \cdot (r - s) > 0$ and hence $t < 0$, and similarly $s < 0$. But this forces both x and y to be negative. □

Proposition 4.7.10

(i) *In both R_d and R_m, we have*

$$((\exists y\colon R)(x \cdot y = 1) \dashv\vdash_x ((x > 0) \vee (x < 0))) .$$

(ii) *R_d and R_m are both residue fields.*

(iii) *R_d is a local ring.*

Proof (i) One direction follows easily from 4.7.9. Conversely, suppose $x > 0$; then we may define its multiplicative inverse to be the section

$$\langle \{q\colon Q \mid ((q \leq 0) \vee ((q > 0) \wedge (1/q \in U_x)))\}, \{q\colon Q \mid ((q > 0) \wedge (1/q \in L_x))\}\rangle$$

(note that $x > 0$ implies $1/n \in \ulcorner L_x \urcorner$ for some natural number n, so that the upper section of this pair is inhabited). It is again straightforward to verify that this (is a section of the required type and) is a multiplicative inverse for x. Similarly if $x < 0$.

(ii) follows immediately from (i) and 4.7.6.

(iii) In R_d, we have $(\forall x)((x > 0) \vee (x < 1))$, and hence by (i) we have $(\forall x)((\exists y)(xy = 1) \vee (\exists y)((1 - x)y = 1))$. But this is equivalent to the local-ring axiom as stated in 1.1.7(g). □

However, R_m is not a local ring in general (see 4.7.11 below). On the other hand, we can say a good deal more about the algebraic character of R_d: it is a 'separably real-closed local ring'. This means that it satisfies the axioms of a coherent theory (which we shall not state here explicitly) whose models in **Set** are Henselian local rings with real-closed residue fields. (Once again, we note that the stalks of the sheaf of continuous real-valued functions on any topological space are well known to be rings of this type.) For details of the proof, see [611, 511].

Now we revert to the question, which we deferred earlier, of when the objects R_d and R_m coincide. The answer turns out to be straightforward:

Theorem 4.7.11 *In a topos $\mathcal{E}$ with natural number object, the following conditions are equivalent:*

(i) *$\mathcal{E}$ satisfies De Morgan's law (cf.* 4.6.2*).*

(ii) *The objects R_d and R_m coincide in $\mathcal{E}$.*

(iii) *R_m is a local ring.*

(iv) *The relation $(x < y) \vee (y < x)$ defines an apartness relation on R_m.*

(v) *R_d is conditionally order-complete (cf.* 4.7.7*).*

Proof First suppose De Morgan's law holds, and let x be a MacNeille real in $\mathcal{E}$. To show it is a Dedekind real, suppose given rationals p, q with $p < q$; let $r = (p + q)/2$. By disjointness of L_x and U_x, plus De Morgan's law, we have $(\neg(r \in U_x) \vee \neg(r \in L_x))$; but since $p < r$ the first disjunct implies $(p \in L_x)$, and the second similarly implies $(q \in U_x)$. Thus (i) implies (ii); and the fact that (ii) implies (iii), (iv) and (v) is clear from the results already established about R_d and R_m.

In the converse direction, the key idea is that if we are given a truth-value $w\colon \Omega$, we may construct a MacNeille real $x(w)$ by setting

$$x(w) = \langle \{q\colon Q \mid ((q < 0) \vee ((q < 1) \wedge \neg\neg w))\}, \{q\colon Q \mid ((q > 1) \vee ((q > 0) \wedge \neg w))\}\rangle .$$

It is an easy exercise to verify that $x(w)$ is indeed a MacNeille real. However, if it were a Dedekind real then we should have $((0 \in L_{x(w)}) \vee (1 \in U_{x(w)}))$, which is equivalent to $(\neg\neg w \vee \neg w)$; since w was arbitrary, this shows that De Morgan's law follows from (ii). Also, if the indicated relation were an apartness relation on R_m, then we should have $((0 < x(w)) \vee (x(w) < 1))$ (since we certainly have $(0 \leq x(w) \leq 1)$), but this again reduces to the statement of De Morgan's law. Similarly, if R_m is assumed to be a local ring, then either $x(w)$ or $1 - x(w)$ must be invertible, which yields the same conclusion via 4.7.10(i). Finally, we note that the set

$$S_w = \{q\colon Q \mid ((q = 0) \vee ((q = 1) \wedge w))\}$$

is inhabited and bounded in R_d. But if this set has a supremum ($y(w)$, say) in R_d, then we have $(w \vdash (y(w) = 1))$ and $(\neg w \vdash (y(w) = 0))$, and so from $((0 < y(w)) \vee (y(w) < 1))$ we may deduce $(\neg\neg w \vee \neg w)$. Thus, once again, we obtain De Morgan's law. □

We remark that the elements $x(w)$ constructed in the proof of 4.7.11 are all idempotents in the ring R_m; for they satisfy $\neg\neg((x(w) = 0) \vee (x(w) = 1))$ and hence $\neg\neg(x(w)^2 = x(w))$, from which the result follows by $\neg\neg$-separatedness of R_m. In fact the Boolean algebra of idempotents of R_m is readily seen to be isomorphic to $\Omega_{\neg\neg}$ (whereas, since R_d is a local ring, its algebra of idempotents is simply the two-element set $\{0, 1\}$). This fact may be used in conjunction with 4.6.2(vii) to give another proof that the equality $R_d = R_m$ implies De Morgan's law. Also, the failure of R_m to be an indecomposable ring, in general, may be exploited to give another construction of the Gleason cover $\gamma\mathcal{E}$ described in the previous section: in fact $(\gamma\mathcal{E}, R_d)$ is the Pierce representation of the ringed topos $(\mathcal{E}, R_m)$ (cf. B4.4.23). For more details, see [512].

Whilst we are on the subject of comparisons between different objects of real numbers, it seems appropriate to mention the other popular way of defining the reals from the rationals, via equivalence classes of Cauchy sequences. This leads to an object commonly called the *Cauchy real number object* (though 'Cantor real number object' would be more historically accurate; fortunately our notation R_c will do for either.) When working constructively with Cauchy sequences, it is convenient to restrict attention to sequences that converge 'at a uniform rate'; thus we define the object $C \subseteq Q^N$ of Cauchy sequences to be the interpretation of the term

$$\{f\colon Q^N \mid (\forall m, n\colon N)((0 < m \leq n) \Rightarrow (|f(m) - f(n)| < 1/m))\}\,.$$

With this definition, we can then define the equivalence relation $E \subseteq C \times C$ of 'converging to the same limit' by the term

$$\{\langle f, g\rangle\colon C \times C \mid (\forall n\colon N)((0 < n) \Rightarrow (|f(n) - g(n)| < 3/n))\}\,.$$

It requires a little work to verify that E is indeed an equivalence relation (specifically, that it is transitive); but if we assume $\langle f, g\rangle \in E$, $\langle g, h\rangle \in E$ and $n > 0$.

we have $|f(n) - f(6n)| < 1/n$, $|f(6n) - g(6n)| < 1/2n$, $|g(6n) - h(6n)| < 1/2n$ and $|h(6n) - h(n)| < 1/n$, yielding the required inequality. The object R_c is then defined to be the quotient of C by E; we may proceed to define the arithmetic operations and order relations on C (for example, we define addition by $f + g = \lambda n . f(2n) + g(2n)$ to ensure that the sequence converges at the required rate), and check that they induce the appropriate structure on the quotient. We may also define a mapping $h: C \to R_d$ by the term

$$\langle \{q: Q \mid (\exists n)(q < f(n) - 1/n)\}, \{q: Q \mid (\exists n)(q > f(n) + 1/n)\} \rangle$$

where f is a free variable of type C; again, it requires verification that this is indeed a Dedekind real, but if we are given $p < q$ in Q then we can choose n with $3/n < q - p$, and then we must have either $p < f(n) - 1/n$ or $q > f(n) + 1/n$. It is similarly straightforward to verify that we have

$$(\forall f, g: C)((h(f) = h(g)) \Leftrightarrow (\langle f, g \rangle \in E)) ,$$

so that h induces a monomorphism $R_c \to R_d$. Moreover, this monomorphism preserves the algebraic and order structure.

However, $R_c \rightarrowtail R_d$ is far from being an isomorphism in general. The reason can be summarized in two words: 'countable choice'. If we are in a context where the axiom of countable choice holds (in other words, a topos where the natural number object is projective), then we may express any Dedekind real as the limit of a Cauchy sequence of rationals, since for every Dedekind real x and every nonzero natural number n we may find a rational q with $|x - q| < 1/n$. (To do this, note first that since L_x and U_x are inhabited we may certainly find rationals p, r with $p < x < r$; then we may divide the interval $[p, r]$ into finitely many subintervals of length at most $1/n$.) But in a topos where the axiom is not valid, R_c may be very much smaller than R_d. We give two examples to show the different types of behaviour that can occur:

Examples 4.7.12 (*a*) Let $\mathcal{E}$ be the topos of sheaves on a locally connected locale X (or more generally any locally connected topos over **Set**). Then the inverse image functor $\gamma^*: \mathbf{Set} \to \mathcal{E}$ preserves exponentials (cf. C1.5.9), and so $Q^N \cong \gamma^*(\mathbb{Q}^{\mathbb{N}})$ is the sheaf whose sections over an open $U \subseteq X$ are simply the locally constant functions $U \to \mathbb{Q}^{\mathbb{N}}$. Similarly, C is the sheaf whose sections over U are the locally constant Cauchy-sequence-valued functions on U, and E is also a constant sheaf; so we deduce that R_c is simply the sheaf $\gamma^*(\mathbb{R})$ of locally constant (Cauchy-)real-valued functions over X. In contrast, R_d may be substantially larger than $\gamma^*(\mathbb{R})$, for example if X itself is $\mathbb{R}$.

(*b*) Let $\mathcal{E}$ be a localic **Set**-topos satisfying (SS), for example the topos of sheaves on a second countable zero-dimensional locale X. Then we saw in 4.5.16 that we may 'lift' the axiom of dependent choices from **Set** to $\mathcal{E}$, and hence prove internally that every Dedekind real in $\mathcal{E}$ is the limit of a Cauchy sequence of rationals; thus in such a topos we have $R_c \cong R_d$. If $\mathcal{E}$ is the topos of 4.5.2(*a*), this

result is not surprising, since the corresponding locale has the property that every continuous real-valued function on it is constant; but there are other examples provided by 4.5.2(d), for example the topos $\mathbf{Sh}(\mathbb{Q})$ (where $\mathbb{Q}$ is topologized as a subspace of $\mathbb{R}$), in which there are many continuous real-valued functions which are not locally constant, and so R_c is very far from coinciding with $\gamma^*(\mathbb{R})$ in this case.

It can be shown that, for a locale X, R_c coincides with $\gamma^*(\mathbb{R})$ in $\mathbf{Sh}(X)$ iff, for every open $U \subseteq X$, the lattice of (relatively) clopen sublocales of U is closed under countable unions. (For a second countable locale X, this condition is equivalent to local connectedness.) For details of the proof, see [357].

The fact that, in a spatial topos, R_c need not coincide with either the sheaf of continuous real-valued functions or the sheaf of locally constant real-valued functions means that it is generally regarded as less useful than R_d as a model of 'what the real numbers ought to be', even though its algebraic and order structure is in general rather easier to work with. Another reason for discarding R_c is that, when we regard it as an internal topological space in $\mathcal{E}$ (equipped with the topology generated by rational open intervals, i.e. the image in $P(R_c)$ of the frame $\mathcal{F}(R)$ of formal reals defined in 4.7.4), it is not in general sober (cf. C1.6.8); whereas R_d, being by definition the space of points of the formal reals, is always sober.

Example 4.7.13 This seems a good point at which to give our promised example of a topos $\mathcal{E}$ in which the Heine–Borel theorem fails, so that the locale R_f of formal reals does not coincide with the space R_d. The first such example was given by M. Fourman and M. Hyland [364]; we present a slightly different example due to A. Joyal, which constructs $\mathcal{E}$ as the classifying topos of a geometric propositional theory, namely the 'theory of an open cover $\mathcal{U}$ of $[0,1]$ with no finite subcover'. Actually, in order to say 'with no finite subcover' in a positive (geometric) way, we demand a stronger condition: namely, that each finite subset of $\mathcal{U}$ should have total Lebesgue measure strictly less than $\frac{1}{2}$.

The primitive propositions of our theory $\mathbb{T}$ are formal expressions of the form $((p,q) \in \mathcal{U})$, where p and q are rationals with $0^- \leq p < 1$, $0 < q \leq 1^+$ and $0 < q - p < \frac{1}{2}$. (Note: for the sake of uniformity of notation, we write $(0^-, q)$ for the half-open interval $[0,q) = \{x \in \mathbb{R} \mid 0 \leq x < q\}$, and similarly for $(p, 1^+)$.) For technical reasons, we wish our covering to be a lower set in the poset of rational open intervals, so we adopt the axioms

$$(((p,q) \in \mathcal{U}) \vdash ((p',q') \in \mathcal{U}))$$

whenever $p \leq p' < q' \leq q$. Next, we adopt the axioms

$$\Big(\bigwedge_{i=1}^{n} ((p_i, q_i) \in \mathcal{U}) \vdash \bot\Big)$$

for each finite sequence $((p_1, q_1), \ldots, (p_n, q_n))$ such that the Lebesgue measure of $\bigcup_{i=1}^n (p_i, q_i) \subseteq \mathbb{R}$ is greater than or equal to $\frac{1}{2}$. Finally, for each real number x with $0 \leq x \leq 1$, we adopt the axiom

$$(\top \vdash \bigvee\{((p,q) \in \mathcal{U}) \mid p < x < q\}) .$$

Clearly, $\mathbb{T}$ has no models in **Set**, so its classifying topos will be of the form $\mathbf{Sh}(X)$ where X is a locale with no points; nevertheless, we shall see shortly that it is consistent. We shall write $\mathbb{T}_0$ for the subtheory of $\mathbb{T}$ defined by the first two groups of axioms, and $\mathbf{Sh}(Y)$ for the classifying topos of $\mathbb{T}_0$; much of the argument which follows is concerned with determining the relationship between Y and its sublocale X.

The frame $\mathcal{O}(Y)$ is easy to describe. Consider the poset P whose elements (which we shall call *nodes*) are formal finite conjunctions $\bigwedge_{i=1}^n ((p_i, q_i) \in \mathcal{U})$, subject to the condition that $\bigcup_{i=1}^n (p_i, q_i)$ has Lebesgue measure less than $\frac{1}{2}$ – we refer to this condition by saying that the conjunction is *consistent* – and ordered by setting

$$\bigwedge_{j=1}^m ((r_j, s_j) \in \mathcal{U}) \leq \bigwedge_{i=1}^n ((p_i, q_i) \in \mathcal{U})$$

iff, for each i, there exists j such that $r_j \leq p_i < q_i \leq s_j$. (Thus we move downwards in the poset P by either increasing the number of terms in our conjunction, or enlarging the intervals to which they refer.) Then $\mathcal{O}(Y)$ is simply the frame of lower sets of P; equivalently, $\mathbf{Sh}(Y)$ is simply the functor category $[P^{\mathrm{op}}, \mathbf{Set}]$. Thus $\mathbf{Sh}(X)$ may be identified with $\mathbf{Sh}(P, T)$ for a suitable coverage T on P; equivalently, $\mathcal{O}(X)$ is the frame of T-ideals of P, in the sense of C1.1.16(e). Again, T is easy to describe: for each real number $x \in [0, 1]$, each node n is covered by the set of all conjunctions $n \wedge ((p,q) \in \mathcal{U})$ where $p < x < q$, and the extra interval (p, q) is small enough for this new conjunction to be consistent. Since the definition of consistency involves a strict inequality, it is clear that we can always find such an interval, for any node n and any $x \in [0, 1]$; hence every cover in T is inhabited – equivalently, X is a dense sublocale of Y. In particular, this shows that X is nontrivial; so $\mathbb{T}$ is consistent, as we claimed earlier.

But in fact X satisfies a much stronger condition than denseness in Y; it is actually a flat sublocale of Y as defined in C1.1.16(d), i.e. $\mathcal{O}(X)$ is closed under finite joins in $\mathcal{O}(Y)$ (which are simply unions, of course). To show this, it suffices by C1.1.16(e) to show that the covers in T are (upwards) directed; but this is easy, since if R is a cover of n and we have $n \wedge ((p_i, q_i) \in \mathcal{U}) \in R$ for $i = 1, 2$, then the node $n \wedge ((p_3, q_3) \in \mathcal{U})$, where $p_3 = \max\{p_1, p_2\}$ and $q_3 = \min\{q_1, q_2\}$, is an upper bound for them which lies in R (it is clearly consistent, since the interval (p_3, q_3) is contained in both (p_1, q_1) and (p_2, q_2)).

Now we wish to identify the object R_d in $\mathbf{Sh}(X)$ (or at least the closed interval $[0, 1]_d \rightarrowtail R_d$ which is the interpretation of $\{x : R_d \mid 0 \leq x \leq 1\}$). To do this, we first consider its global sections, that is the actual Dedekind real

numbers in $\mathbf{Sh}(X)$. We note that Y is spatial, and that it has a focal point in the sense of C1.5.6 (that is, a point whose only neighbourhood is the whole space), corresponding to the empty family of rational intervals. So the reflection of Y in the category of Hausdorff spaces is a singleton, and thus Dedekind real numbers in $\mathbf{Sh}(Y)$ are just constant real-valued functions on Y. We now appeal to Joyal's lemma C4.1.15(ii): since the locale corresponding to the frame $\mathcal{F}([0,1])$ is compact Hausdorff, each morphism from X to it can be uniquely extended to a morphism from Y. Thus the Dedekind reals in $\mathbf{Sh}(X)$ which lie between 0 and 1 are simply the constant real numbers in this interval. A similar argument with X and Y replaced by their open sublocales corresponding to a principal ideal $\downarrow(n)$ in P (note that every principal ideal is a T-ideal) yields the same conclusion about the sections of $[0,1]_d$ over any such open in X: that is, $[0,1]_d$, considered as a T-sheaf on P, is no more than the constant functor with value $[0,1]$ in $\mathbf{Set}$. (Again, it is not hard to verify directly that this functor is a T-sheaf; but of course that on its own would not suffice to identify it with the object $[0,1]_d$ as constructed in $\mathbf{Sh}(X)$.) So the generic $\mathbb{T}$-model in $\mathbf{Sh}(X)$ really does define an open cover of the closed interval $[0,1]_d$ in the sense of this topos; and it clearly has no finite subcover.

Remark 4.7.14 Even if R_d is not the whole of R_f, it may easily be seen to be strongly dense in the latter, in the sense introduced in C1.1.22: whatever the extra covers are that we have to add to the presentation of $\mathcal{F}(R)$ to obtain $\mathcal{O}(R_d)$, they are certainly inhabited, since each rational open interval (p,q) contains points. Also, R_d is an open locale in the sense of C3.1.16, because it is spatial. It follows from the 'closed subgroup theorem' C5.3.2 that it cannot be a localic subgroup of R_f unless it coincides with R_f: that is, the addition operation for Dedekind reals, which we (implicitly) defined on the spatial product $(R_d \times R_d)_p$ (we didn't actually verify that it was continuous, but that is easy), cannot be extended to the locale product. (On the other hand, if R_d is locally compact, then the spatial and localic versions of $R_d \times R_d$ coincide, by C4.1.8, and so R_d is a localic group; this provides another proof that the Heine–Borel theorem is sufficient for the isomorphism $R_d \cong R_f$.) This is another reason for preferring R_f to R_d in contexts (such as that just given) where the two differ.

Finally in this section, we describe a universal characterization of the Dedekind reals (or rather, of the closed unit interval therein), due to P. Freyd. We consider posets P equipped with top and bottom elements 1 and 0 which are *DM-separated* in the sense that we have

$$(\forall p\colon P)(\neg(0=p) \vee \neg(p=1)) \ ;$$

the reason for the name is that, if De Morgan's law holds, this is equivalent to the simpler condition $\neg(0=1)$, but in general it is strictly stronger. In particular, we note that this condition holds for the subobject classifier Ω iff De Morgan's law holds in our topos $\mathcal{E}$. (It does, however, hold for the Dedekind unit interval

$[0,1]_d \subseteq R_d$ in any topos, since we have $(\forall x\colon R_d)((0<x)\vee(x<1))$.) We write $\mathbf{dms}(\mathcal{E})$ for the category of posets in $\mathcal{E}$ satisfying this condition (the morphisms being maps preserving order and the two distinguished elements).

Classically, the ordinal sum or 'ordered wedge' of two posets P, Q with top and bottom elements is defined to be the poset obtained from the disjoint union of P and Q by identifying the top element of P with the bottom element of Q. It is well known that it can also be identified with a sub-poset of $P \times Q$, namely

$$\{\langle p,q\rangle : P\times Q \mid ((p=1)\vee(0=q))\}\ .$$

In our constructive context, we have to modify this definition slightly: we define the 'thick ordered wedge' $P \oplus Q$ to be

$$\{\langle p,q\rangle : P\times Q \mid (\neg(p=1)\Rightarrow(0=q))\wedge(\neg(0=q)\Rightarrow(p=1))\}\ .$$

Clearly, each of the two conditions in the definition of $P\oplus Q$ is implied by $((p=1)\vee(0=q))$, so the 'classical ordered wedge' is contained in this one: we can think of it as obtained from the classical ordered wedge by 'thickening up' the latter around the point $\langle 1,0\rangle$ to make the two parts fit together more tightly. Note also that, if P and Q are equipped with apartness relations $\#$ such that equality is the negation of apartness (for example, if they are closed intervals in R_d), then each of the conditions in the definition of $P\oplus Q$ is equivalent to

$$\neg((p\,\#\,1)\wedge(0\,\#\,q))\ ,$$

since the latter is clearly equivalent to $((p\,\#\,1)\Rightarrow(0=q))$ and hence (since the truth-value of $(0=q)$ is $\neg\neg$-closed) to $(\neg\neg(p\,\#\,1)\Rightarrow(0=q))$. However, in general neither of the two conditions in the definition implies the other.

Since the top and bottom elements of $P\oplus Q$ are $\langle 1,1\rangle$ and $\langle 0,0\rangle$, it is clear that it is a DM-separated poset if either P or Q is DM-separated, and hence that the mapping $\langle P,Q\rangle \mapsto P\oplus Q$ defines a bifunctor $\mathbf{dms}(\mathcal{E})\times\mathbf{dms}(\mathcal{E})\to\mathbf{dms}(\mathcal{E})$. We shall be concerned with the 'diagonalization' of this bifunctor, i.e. the functor $P\mapsto P\oplus P$.

Freyd's characterization asserts that $[0,1]_d$ is a terminal coalgebra for this endofunctor of $\mathbf{dms}(\mathcal{E})$. In order to prove this, we first need to show that it is a coalgebra (for which the structure map $[0,1]_d \to [0,1]_d\oplus[0,1]_d$ is an isomorphism, cf. A1.1.4). But since any closed interval in R_d is isomorphic to any other, it will suffice to prove

Lemma 4.7.15 *For any three Dedekind reals a,b,c with $a<b<c$, we have $[a,c]\cong[a,b]\oplus[b,c]$.*

This lemma is, of course, the reason why we had to adopt the 'thickened' definition of the ordered wedge; it fails to hold for the classical definition.

Proof For convenience, we shall take a,b and c to be $-1,0$ and 1 respectively. The maps establishing the isomorphism are then easy to define: they are

$(x \mapsto \langle \min\{x,0\}, \max\{x,0\}\rangle)$ and $(\langle y,z\rangle \mapsto y+z)$. We have to show that the first map does indeed take values in $[-1,0] \oplus [0,1]$, and that the two maps are inverse to each other.

The first statement is easy to verify, since R_d has an apartness relation: it is enough to verify

$$(\forall x\colon [-1,1])\neg((x<0) \wedge (0<x))\,.$$

But this is immediate from the definition of $<$. The identity $x = \min\{x,0\} + \max\{x,0\}$ holds for all x, not just for those in $[-1,1]$. And if $y \leq 0 \leq z$, then we certainly have $y \leq y+z$ and hence $y \leq \min\{y+z,0\}$; but if we had $y < \min\{y+z,0\}$, then we should have $y<0$, and also $y < y+z$ and hence $0<z$, which contradicts the assumption that the pair $\langle y,z\rangle$ belongs to $[-1,0] \oplus [0,1]$. Thus we have $y \geq \min\{y+z,0\}$, and hence $y = \min\{y+z,0\}$. Similarly, $z = \max\{y+z,0\}$ whenever $\langle y,z\rangle \in [-1,0] \oplus [0,1]$. □

For future reference, we note that the coalgebra structure on $[0,1]$ which we obtain from 4.7.15 is given by the mapping

$$x \mapsto \langle \min\{2x,1\}, \max\{2x-1,0\}\rangle \colon [0,1] \to [0,1] \oplus [0,1]$$

(and its inverse is given by 'averaging': $\langle y,z\rangle \mapsto (y+z)/2$).

Now consider an arbitrary coalgebra $(c\colon X \to X \oplus X)$. It will be convenient to consider not c itself but its two coordinate projections $d,u\colon X \rightrightarrows X$: note that these maps are order-preserving and satisfy

$$(\forall x\colon X)(\neg(dx=1) \Rightarrow (0=ux))$$

and the dual condition. We first need to verify

$$(\forall x\colon X)((0=ux) \vee ((0=uux) \wedge (ddx=1)) \vee (dx=1)) \qquad (*)$$

For this, observe that for any x we have $(\neg(0=dux) \vee \neg(dux=1))$, since X is DM-separated; from $\neg(dux=1)$ we may deduce $(0=uux)$, and from $\neg(0=dux)$ we may deduce $\neg(0=ux)$ (since d preserves 0) and hence $(dx=1)$. Combining this with the similar deduction from $(\neg(0=udx) \vee \neg(udx=1))$ yields the displayed formula above.

To define a coalgebra homomorphism $f\colon X \to [0,1]_d$, we suppose given an element x of X, and proceed to construct a Dedekind real $\langle L_x, U_x\rangle$ as follows. We consider the binary tree T whose nodes are all formal finite strings of d's and u's (including the empty string $[\,]$), ordered by $v \leq w$ iff $w = sv$ for some string s; given such a string w, we write wx for the element of X that we get by evaluating this string at x. We also associate with each node w a dyadic rational r_w (that is, one whose denominator is a power of 2) in $(0,1)$, in a straightforward way; namely $r_{[\,]} = \frac{1}{2}$ and, if w is a string of length $n \geq 0$, then $r_{dw} = r_w - 1/2^{n+2}$ and $r_{uw} = r_w + 1/2^{n+2}$. (Equivalently, r_w is the unique real number mapped

by w to $\frac{1}{2}$ in the coalgebra structure already described on $[0,1]$.) Clearly, every dyadic rational in $(0,1)$ occurs as r_w for a unique node w.

We now construct $\langle L_x, U_x\rangle$ by the following recursive procedure. Initially L_x contains all rationals less than 0, and U_x all rationals greater than 1. We work up the binary tree examining each node in turn; by 'examining a node' w, we mean asking the two questions 'Is $0 = uwx$?' and 'Is $dwx = 1$?'. If $0 = uwx$, we put all rationals greater than r_w into the set U_x, and if $dwx = 1$ we put all rationals less than r_w into L_x. Formally, we have

$$L_x = \{q\colon Q \mid (q < 0) \vee (\exists w\colon T)((dwx = 1) \wedge (q < r_w))\}$$

with a similar definition for U_x.

Lemma 4.7.16 *The mapping $x \mapsto \langle L_x, U_x\rangle$ defined above is a coalgebra homomorphism $f\colon X \to [0,1]_d$.*

Proof First we have to verify that $\langle L_x, U_x\rangle$ really is a Dedekind real (it clearly lies in $[0,1]$). It is also clear that L_x is an open lower section and U_x is an open upper section, so we have to verify that they are disjoint and satisfy the 'zero distance apart' condition. For disjointness, we have to show that if $r_v < r_w$ (for two nodes v, w of T), then we cannot have both $0 = uvx$ and $dwx = 1$. It is clear that this cannot occur if v is an ancestor of w (that is, if $w = sv$ for some string s); for then the condition $r_v < r_w$ would imply that uv was also an ancestor of w and so $0 = uvx$ would imply $0 = wx$ and hence $0 = dwx$. Similarly if w is an ancestor of v. So suppose neither is an ancestor of the other, and let z be their longest common ancestor; again since $r_v < r_w$, v must have the form sdz and w must have the form tuz for some strings s, t. Now we have $(\neg(0 = uvx) \vee \neg(uvx = 1))$; but from $\neg(uvx = 1)$ we can deduce $\neg(dzx = 1)$ since the composite us preserves 1. Similarly, from $\neg(0 = dwx)$ we could deduce $\neg(0 = uzx)$; but we cannot have both of these since the pair $\langle dzx, uzx\rangle$ belongs to $X \oplus X$. So we must have either $\neg(0 = uvx)$ or $\neg(dwx = 1)$.

For the 'zero distance apart' condition, we use the displayed formula $(*)$ above: the point of this formula is that it assures us that, by the time we have examined all nodes of length at most n, we shall have 'tied down' fx to lie in an interval of length at most $1/2^n$. Thus, given rationals q and r with $q < r$, we choose n such that $1/2^n < r - q$, and then by examining all nodes of length at most n we shall succeed in establishing either $(q \in L_x)$ or $(r \in U_x)$.

Next, we must verify that f is order-preserving. But this follows easily from the fact that d and u are order-preserving: if $x_1 \leq x_2$ in X, then for any word $w \in T$ we have $wx_1 \leq wx_2$, and so $wx_1 = 1$ implies $wx_2 = 1$; thus we obtain $L_{x_1} \subseteq L_{x_2}$. A similar argument shows that f preserves 0 and 1.

Finally, we must show that f is a coalgebra homomorphism. But it is easy to verify that for any node w of T we have $r_{wd} = r_w/2$, and hence that $f(dx) = \min\{2f(x), 1\}$, as required (and similarly $f(ux) = \max\{2f(x) - 1, 0\}$). □

Theorem 4.7.17 *In any topos $\mathcal{E}$, the Dedekind unit interval $[0,1]_d$ is a terminal coalgebra for the endofunctor $X \mapsto X \oplus X$ of* $\mathbf{dms}(\mathcal{E})$.

Proof After the last two lemmas, we have only to verify that the coalgebra homomorphism f constructed above is unique. Suppose we have another such homomorphism g, and suppose we have an $x\colon X$ such that $(fx < gx)$. Choose a dyadic rational r such that $fx < r < gx$, and let w be the unique node of T such that $r_w = r$. Then we have $0 = uwx$ since $r \in U_x$; however, since g commutes with u and d, it is easy to see that we have $0 < g(uwx)$, since $r = r_w$ is the largest element of $[0,1]$ such that $uwr = 0$. Thus we have contradicted our supposition; that is, we have proved

$$(\forall x\colon X)\neg(fx < gx)\ .$$

Similarly, we may prove $(\forall x\colon X)\neg(gx < fx)$; so we conclude $(\forall x\colon X)(fx = gx)$, i.e. $f = g$. □

Remarks 4.7.18 (*a*) It will be observed that the order-relation on the posets in $\mathbf{dms}(\mathcal{E})$ does not play any essential rôle in the foregoing argument. Thus we could alternatively work in a category whose objects are simply objects of $\mathcal{E}$ equipped with two distinguished elements which are DM-separated, and describe $[0,1]_d$ as the terminal coalgebra of the 'thick wedge' endofunctor of this category.

(*b*) It is also interesting to note that, if we weaken the DM-separatedness condition to mere separatedness (that is, to the condition $\neg(0 = 1)$), then the terminal coalgebra for the corresponding endofunctor of our enlarged category is simply the MacNeille unit interval $[0,1]_m$. Once again, most of the preceding proof goes through in this case without significant modification.

Suggestions for further reading: Banaschewski & Mulvey [61, 62], Dubuc [301], Fourman [357], Fourman & Hyland [364], Kennison & Ledbetter [598], Negri & Soravia [890], Reichman [1001], Rousseau [1071], Stout [1123].

D5

ASPECTS OF FINITENESS

D5.1 Natural number objects revisited

The definition and some of the basic properties of natural number objects in a topos were introduced in Section A2.5. However, there were some important properties which could not be developed there because they required the ability to interpret higher-order logic in a topos. Having developed the techniques for this in the previous chapter, we are now ready to return to the study of natural number objects and complete what we left unfinished in Section A2.5.

Let us, for the time being, write $\mathbb{T}$ for the algebraic theory freely generated by one nullary and one unary operation (so that a $\mathbb{T}$-algebra is simply an object A equipped with morphisms $a\colon 1 \to A$ and $t\colon A \to A$). We recall that, if (N, o, s) is a natural number object in a topos $\mathcal{E}$ (that is, an initial object in the category of $\mathbb{T}$-algebras in $\mathcal{E}$), then we have two colimit diagrams $1 \to N \leftarrow N$ and $N \rightrightarrows N \to 1$ (A2.5.5), and N satisfies the 'Peano postulates' that s is monic, o and s are disjoint subobjects of N, and any (o, s)-closed subobject of N is the whole of N (A2.5.8 and A2.5.9). Our first objective in this section is to prove that the converses of these results are valid. For the moment, we shall call a $\mathbb{T}$-algebra (N, o, s) a *Lawvere natural number object* if it satisfies Definition A2.5.1, a *Freyd natural number object* if the two diagrams of A2.5.5 are colimits, and a *Peano natural number object* if it satisfies the Peano postulates.

Given a $\mathbb{T}$-algebra (A, a, t), we shall say a subobject $A' \rightarrowtail A$ is *closed* (or (a, t)-*closed*, if it is necessary to specify the structure maps) if it is a sub-$\mathbb{T}$-algebra of A, i.e. if there exist morphisms a' and t' making

$$\begin{tikzcd}
1 \arrow[r, "a'"] \arrow[dr, "a"'] & A' \arrow[d, tail] & A' \arrow[l, "t'"'] \arrow[d, tail] \\
 & A & A \arrow[l, "t"]
\end{tikzcd}$$

commute. And we shall say that A' is *recursive* if, in addition, the pair (a', t') is jointly epic.

Lemma 5.1.1 *Any $\mathbb{T}$-algebra contains a unique smallest closed subobject, which is recursive.*

Proof It is easy to see that an arbitrary (internal) intersection of sub-$\mathbb{T}$-algebras is a sub-$\mathbb{T}$-algebra (cf. 5.3.2 below), so we may obtain the smallest one by forming the intersection of all of them. Also, if (A, a, t) is a $\mathbb{T}$-algebra, the subobject which is the union of the images of a and t is (a, t)-closed, so if A is minimal in the above sense this subobject must be the whole of A. □

Theorem 5.1.2 *For a $\mathbb{T}$-algebra (N, o, s) in a topos $\mathcal{E}$, the following are equivalent:*

(i) *(N, o, s) is a Lawvere natural number object.*

(ii) *(N, o, s) is a Freyd natural number object.*

(iii) *(N, o, s) is a Peano natural number object.*

Proof (i) ⇒ (ii) and (i) ⇒ (iii) were proved in Section A2.5.

(ii) ⇒ (iii): Clearly, the first of Freyd's two conditions implies that s is monic and its image is disjoint from o. So we have only to verify the induction postulate. Suppose $m: N' \rightarrowtail N$ is a (o, s)-closed subobject of N; by 5.1.1, we might as well assume that N' is actually recursive. Now consider the binary relation $S \rightarrowtail N \times N$ which is the symmetrization of $(1, s): N \rightarrowtail N \times N$, i.e. the union of this relation and its opposite. The assertion that N' is (o, s)-recursive implies that the formula

$$(\forall x, y : N)((\langle x, y\rangle \in \ulcorner S \urcorner) \Rightarrow ((x \in \ulcorner N' \urcorner) \Leftrightarrow (y \in \ulcorner N' \urcorner)))$$

is valid. Now let $T \rightarrowtail N \times N$ be the interpretation of the closed term

$$\{z : N \times N \mid (\forall w : PN)((w \text{ is } S\text{-closed}) \Rightarrow ((\mathsf{fst}(z) \in w) \Leftrightarrow (\mathsf{snd}(z) \in w)))\},$$

where 'w is S-closed' is an abbreviation for the formula

$$(\forall x, y)\,((\langle x, y\rangle \in \ulcorner S \urcorner) \Rightarrow ((x \in w) \Leftrightarrow (y \in w)))\ .$$

It is clear that T is an equivalence relation, and that it contains S; so it must contain the kernel-pair of the coequalizer of $S \rightrightarrows N$. But this coequalizer is $N \to 1$, since S contains the graph of s; so T is the whole of $N \times N$. Also, since N' is S-closed, it is also T-closed, i.e.

$$(\forall x, y : N)((\langle x, y\rangle \in \ulcorner T \urcorner) \Rightarrow ((x \in \ulcorner N' \urcorner) \Leftrightarrow (y \in \ulcorner N' \urcorner)));$$

but since $(o \in \ulcorner N' \urcorner)$ holds, it follows that we have

$$(\forall x : N)(x \in \ulcorner N' \urcorner),$$

i.e. N' is the whole of N.

(iii) $\Rightarrow$ (i): Suppose given a $\mathbb{T}$-algebra (A, a, t). Applying 5.1.1 to $N \times A$, we obtain a diagram

$$\begin{tikzcd}
1 \arrow[r, "b"] \arrow[dr, "{(o,a)}"'] & B \arrow[d, tail, "{(p,q)}"] & B \arrow[l, "u"'] \arrow[d, tail, "{(p,q)}"] \\
& N \times A & N \times A \arrow[l, "s \times t"']
\end{tikzcd}$$

where (b, u) is jointly epic. We shall show that (p, q) is the graph of a morphism $N \to A$, i.e. that p is an isomorphism.

It is clear that p is epic, since its image is a closed subobject of N. To show that it is monic, consider the subobject $N' \rightarrowtail N$ which is the interpretation of the term

$$\{x : N \mid (\forall y, y' : B)(py = x = py') \Rightarrow (y = y')\};$$

we shall show that N' is (o, s)-closed, and so is the whole of N. First, $o \in \ulcorner N' \urcorner$ because the square

$$\begin{tikzcd}
1 \arrow[r, "b"] \arrow[d] & B \arrow[d, "p"] \\
1 \arrow[r, "o"] & N
\end{tikzcd}$$

is not merely commutative but a pullback, since B is the union of b and the image of u, and the latter is mapped by p into the subobject s which is disjoint from o. Similarly, the square

$$\begin{tikzcd}
B' \arrow[r, tail] \arrow[d, "p'"] & B \arrow[d, "p"] \\
N \arrow[r, tail, "s"] & N
\end{tikzcd}$$

is a pullback, where B' is the image of u (note that we don't yet know that u is monic, but it doesn't matter) and p' is the unique factorization of p through $B \twoheadrightarrow B'$ (which exists because s is monic). Thus from $(py = sx = py')$ we may deduce

$$(\exists z, z' : B)((uz = x) \wedge (uz' = x') \wedge (pz = x = pz')),$$

and hence we obtain $(x \in \ulcorner N' \urcorner) \Rightarrow (sx \in \ulcorner N' \urcorner)$.

Thus we have verified that (N, o, s) satisfies the 'existence' half of Definition A2.5.1. For the uniqueness, suppose $f, g : (N, o, s) \rightrightarrows (A, a, t)$ are two

$\mathbb{T}$-algebra homomorphisms; then the equalizer of f and g is a (o, s)-closed subobject of N, and so $f = g$. $\square$

Corollary 5.1.3 *A topos $\mathcal{E}$ has a natural number object iff it contains an object A equipped with a monomorphism $t\colon A \rightarrowtail A$ and a well-supported subobject $B \rightarrowtail A$ disjoint from t.*

Proof First suppose there exists a morphism $a\colon 1 \to A$ factoring through B. Then, applying 5.1.1 to the $\mathbb{T}$-algebra (A, a, t), we obtain a subobject of A which is readily seen to be a Peano natural number object. In general, we may apply this argument to the object $B^*(A)$ of $\mathcal{E}/B$ to prove that the latter topos has a natural number object; but since B is well-supported the canonical geometric morphism $\mathcal{E}/B \to \mathcal{E}$ is a surjection by A4.2.7(a), and so $\mathcal{E}$ has a natural number object by A4.2.1(v). $\square$

For 2-valued toposes (i.e. those in which the only subterminal objects are 1 and 0), we have an even more striking result:

Corollary 5.1.4 *Suppose $\mathcal{E}$ is a 2-valued topos. Then the following conditions are equivalent:*

(i) *$\mathcal{E}$ has a natural number object.*

(ii) *There exists a monic endomorphism $f\colon A \rightarrowtail A$ in $\mathcal{E}$ which is not an isomorphism (i.e. $\mathcal{E}$ contains an object which is not 'Dedekind-finite').*

(iii) *There exists an epic endomorphism $g\colon A \twoheadrightarrow A$ in $\mathcal{E}$ which is not an isomorphism.*

Proof (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are trivial.

(ii) $\Rightarrow$ (i): Given f as in (ii), consider $f^A\colon A^A \rightarrowtail A^A$. The transpose $1 \to A^A$ of 1_A does not factor through f^A, because f is not a split epimorphism, so the pullback of f^A along it is a proper subobject of 1 and hence must be 0. Hence we can apply 5.1.3.

(iii) $\Rightarrow$ (i): Given g as in (iii), we argue similarly using $A^g\colon A^A \rightarrowtail A^A$. $\square$

We remark in passing that Dedekind-finite objects in general toposes have been studied by L. N. Stout [1127].

Remark 5.1.5 There are analogues of 5.1.1, 5.1.2 and 5.1.3 for list objects, as introduced in A2.5.15: given an object B equipped with morphisms $b\colon 1 \to B$ and $t\colon A \times B \to B$ (where A is a given object), there is a unique smallest subobject

$B' \rightarrowtail B$ which is (b,t)-closed in the sense that we have a commutative diagram

$$\begin{tikzcd} 1 \arrow[r, "b'"] \arrow[dr, "b"'] & B' \arrow[d, tail] & A \times B' \arrow[l, "t'"'] \arrow[d, tail] \\ & B & A \times B \arrow[l, "t"] \end{tikzcd} ;$$

for this subobject, the pair (b', t') is jointly epic, and the subobject will have the universal property of a list object LA provided t' is monic and its image is disjoint from b'. The proofs are essentially similar to those given above.

We note that this provides an alternative, less explicit, proof of the existence of list objects in a topos with natural number object (A2.5.17): given an object A, the object $(1 \amalg A)^N$ admits morphisms $(\nu_1)^N : 1 \cong 1^N \to (1 \amalg A)^N$ and

$$A \times (1 \amalg A)^N \overset{\nu_2 \times 1}{\rightarrowtail} (1 \amalg A) \times (1 \amalg A)^N \cong (1 \amalg A)^{(1 \amalg N)} \cong (1 \amalg A)^N$$

whose images are disjoint, so it contains a subobject which is a list object over A.

We saw in Section A2.5 that the natural number object comes equipped with a total ordering $N_1 \rightarrowtail N \times N$. It is natural to ask whether this is in any sense a well-ordering. Since there exists a monomorphism $1 \amalg 1 \rightarrowtail N$, the result of 4.5.7 tells us that it cannot be a well-ordering as defined in Section D4.5, unless the ambient topos $\mathcal{E}$ is Boolean; but, given this impossibility, it comes as close as it can. We say an ordered object $(A, \leq)$ is *decidably well-ordered* if it satisfies the sentence in the internal language which asserts that every inhabited complemented subobject of A has a least element; and we shall see in the next section (5.2.10(ii)) that $(N, \leq)$ has this property in any topos – in particular, it is well-ordered in any Boolean topos. The proof is deferred to Section D5.2, because it involves properties of finite cardinals; however, we note one consequence of it here, which extends a result already observed in Section A2.5.

Proposition 5.1.6 *Let $\mathcal{E}$ be a topos with a natural number object. Then every partial recursive function f of k variables is 'realizable' by a partial map $\overline{f} : N^k \rightharpoonup N$ in the sense that, for standard natural numbers $\overline{n_i} = s^{n_i} o$ in $\mathcal{E}$, the morphism $(\overline{n_1}, \ldots, \overline{n_k}) : 1 \to N^k$ factors through the domain of $\overline{f}$ iff $f(n_1, \ldots, n_k)$ is defined, and we then have*

$$\overline{f}(\overline{n_1}, \ldots, \overline{n_k}) = \overline{f(n_1, \ldots, n_k)} .$$

Proof We use the result that every partial recursive function is constructible from primitive recursive functions plus a single application of the μ-operator: specifically, for any f, there exist primitive recursive functions g

and h (of $k+1$ variables), such that

$$f(n_1,\ldots,n_k) = g(n_1,\ldots,n_k,\mu n\,.\,(h(n_1,\ldots,n_k,n)=0))$$

where '$\mu n\,.\,(\ldots)$' means 'the least n such that ...' (and $f(n_1,\ldots,n_k)$ is undefined if $h(n_1,\ldots,n_k,n) \neq 0$ for all n). We saw in A2.5.2 how to realize primitive recursive functions of $k+1$ variables by morphisms $N^{k+1} \to N$ in $\mathcal{E}$; so we have in particular such a morphism $\overline{h}$. Now the subobject $\overline{h}^*(o)$ is complemented in N^{k+1}, since o is complemented in N; so we may regard it as a complemented subobject of $(N^k)^*(N)$ in $\mathcal{E}/N^k$. Let $m\colon D \rightarrowtail N^k$ be the interpretation of the sentence which asserts that this subobject is inhabited; then in $\mathcal{E}/D$ we have an inhabited complemented subobject of D^*N, so its least element defines a natural number which, when regarded as a partial map $d\colon N^k \rightharpoonup N$, realizes the function $\mu n\,.\,(h(n_1,\ldots,n_k,n)=0)$. (Note that we do *not* assert that D is a complemented subobject of N^k; in general it need not be.) Then we form the partial map

$$\begin{tikzcd} D \arrow[r, "{(m,d)}"] \arrow[d, tail, "m"] & N^{k+1} \arrow[r, "\overline{g}"] & N \\ N^k & & \end{tikzcd}$$

which is the required realization of f. $\square$

The reader should be warned, however, that if f is (classically) a total recursive function $\mathbb{N}^k \to \mathbb{N}$, the partial map $\overline{f}\colon N^k \rightharpoonup N$ constructed in the proof of 5.1.6 need not be a morphism $N^k \to N$ in $\mathcal{E}$ – its domain contains all the standard natural numbers, but the union of the latter need not be the whole of N. (We shall return to this point in Section F3.4.) In this connection, we say a topos $\mathcal{E}$ with a natural number object is *standard* if the standard natural numbers in $\mathcal{E}$ (those of the form $s^n o$ for some external natural number n) form an epimorphic family of morphisms $1 \to N$. We saw in Section A2.5 that in any topos with countable copowers N is the countable copower of copies of 1, with the standard natural numbers as the coprojections; so all such toposes are standard. However, not all toposes with natural number objects are standard:

Example 5.1.7 Let Φ be a filter of subsets of the set $\mathbb{N}$ of natural numbers containing all cofinite sets, and let $\mathcal{E}$ be the filterpower $(\mathbf{Set}/\mathbb{N})_\Phi$ (cf. A2.1.13). Since the projection $P_\Phi\colon \mathbf{Set}/\mathbb{N} \to (\mathbf{Set}/\mathbb{N})_\Phi$ is logical, $P_\Phi(\mathbb{N}^*\mathbb{N})$ is a natural number object in $\mathcal{E}$. It is clear that there are natural numbers in $\mathcal{E}$ which are not standard: for example, the number which is the image under P_Φ of the diagonal map $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$. But P_Φ maps every finite subset of $\mathbb{N}$ to 0, so this natural number is actually disjoint from every standard natural number in $\mathcal{E}$.

Since epimorphic families are stable under pullback in any topos, the standard natural numbers cannot form an epimorphic family in $\mathcal{E}$.

We note that the topos of 5.1.7 is not merely nonstandard; it does not admit any bicartesian (that is, cartesian and cocartesian) functor to any standard topos. For any such functor would preserve the natural number object (and hence the standard natural numbers), by A2.5.6; but it would also preserve equalizers of pairs of morphisms $1 \rightrightarrows N$, and hence map the particular natural number constructed in 5.1.7 to a natural number disjoint from all the standard ones. On the other hand, we have

Proposition 5.1.8 *Let $\mathcal{E}$ be a small standard topos. Then there is a full and faithful bicartesian functor from $\mathcal{E}$ to a Grothendieck topos.*

Proof Let $\mathcal{F}$ be the topos of sheaves on $\mathcal{E}$ for the coverage consisting of all epimorphic families in $\mathcal{E}$. It is easy to see that this coverage is subcanonical, and so the canonical functor $\mathcal{E} \to \mathcal{F}$ is full and faithful (cf. C2.2.15); it preserves finite limits, and all epimorphic families which exist in $\mathcal{E}$ (hence in particular it is a coherent functor). To prove that it preserves arbitrary coequalizers, it therefore suffices to show that it commutes with the construction of the equivalence relation on A generated by a parallel pair $B \rightrightarrows A$ in $\mathcal{E}$; equivalently (since the reflexive and symmetric relation generated by a pair may be constructed using finite limits and images), to show that it commutes with the construction of the transitive closure of a (reflexive and symmetric) relation.

However, in any topos with a natural number object, we may construct the transitive closure as an 'internal N-indexed union'; that is, given a relation $R \rightrightarrows A$, we may construct a relation $R* \rightrightarrows N^*A$ in $\mathcal{E}/N$ such that the pullback of $R*$ along the standard natural number $s^n o$ is the nth iterate $R^{\langle n \rangle}$ of the relation R, and then verify that the image of $\Sigma_N R* \to A \times A$ is the transitive closure $\overline{R}$ of R. (Explicitly, we take $R*$ to be the relation named by the morphism $N \to P(A \times A)$ which is induced by the morphism $1 \to P(A \times A)$ naming the diagonal subobject and the morphism $P(A \times A) \to P(A \times A)$ which internalizes the operation of composing an arbitrary relation $A \looparrowright A$ with R; cf. also 5.5.8 below.) Now since $\mathcal{E}$ is standard, it follows that the family of morphisms $R^{\langle n \rangle} \rightarrowtail \overline{R}$ is epimorphic; so it is preserved by the embedding $\mathcal{E} \to \mathcal{F}$. $\square$

In [371], P. Freyd showed that any topos satisfying the hypotheses of 5.1.8 in fact admits a jointly faithful family of bicartesian functors to **Set**.

The definition of a natural number object tells us that morphisms whose domain is N (or, more generally, an object of the form $A \times N$), may be constructed recursively. However, we frequently wish to construct objects of $\mathcal{E}/N$ recursively, as well: that is, given an object A of $\mathcal{E}$ and a 'procedure' T for constructing new objects from old, we wish to find an object F of $\mathcal{E}/N$ such that $o^*(F) \cong A$ and $s^*(F) \cong T^N(F)$, where T^N denotes the procedure T 'applied fibrewise' to objects of $\mathcal{E}/N$. To assign a precise meaning to the latter informal

idea, we shall assume that T is actually a functor, and indeed an $\mathcal{E}$-indexed functor $\mathbb{E} \to \mathbb{E}$, where $\mathbb{E}$ denotes the canonical indexing of $\mathcal{E}$ over itself as defined in B1.2.2(c). We shall refer to the pair (A, T) as *recursion data* for constructing an object of $\mathcal{E}/N$, and say that F *satisfies* the recursion data (A, T) if $o^*F \cong A$ and $s^*F \cong T^N(F)$. For example, the generic finite cardinal C of A2.5.14 satisfies the recursion data $(0, 1 \amalg (-))$, and the list object $(l\colon LA \to N)$ over A, constructed in A2.5.17, satisfies the recursion data $(1, A \times (-))$.

We shall not consider the general existence question here, since in practice it will always be easy (as it was for the list object) to construct an explicit solution of any recursion problem we encounter. (The general problem is discussed at length in [549], where it is shown that for a general $\mathcal{E}$ there may be recursion problems which do not have a solution, but there is a large class of problems which do.) However, the following general uniqueness result will very frequently be of use to us in the next couple of sections.

Theorem 5.1.9 *Let $\mathcal{E}$ be a topos with a natural number object, A an object of $\mathcal{E}$, and $T\colon \mathbb{E} \to \mathbb{E}$ an $\mathcal{E}$-indexed endofunctor of $\mathbb{E}$. Then there exists, up to canonical isomorphism, at most one object F of $\mathcal{E}/N$ satisfying $o^*(F) \cong A$ and $s^*(F) \cong T^N(F)$.*

Proof Suppose F' is another solution of the recursion problem. We have to construct an isomorphism $F \cong F'$ in $\mathcal{E}/N$; equivalently, we have to construct a morphism $1_N \to \mathrm{Iso}(F, F')$, where $\mathrm{Iso}(F, F')$ denotes the object of isomorphisms constructed in $\mathcal{E}/N$, i.e. the interpretation of the term

$$\{f\colon [F \to F'] \mid (\exists g\colon [F' \to F])((\forall x\colon F)(g(f(x)) = X) \wedge (\forall y\colon F')(f(g(y)) = y))\}$$

in the internal language of this topos. Now pullback functors preserve objects of isomorphisms (since they are logical), so we have a morphism

$$1 \xrightarrow{\overline{1_A}} \mathrm{Iso}(A, A) \cong \mathrm{Iso}(o^*(F), o^*(F')) \cong o^*(\mathrm{Iso}(F, F'))\ .$$

Similarly, since indexed functors are strong by B2.2.2, T induces a morphism

$$\mathrm{Iso}(F, F') \longrightarrow \mathrm{Iso}(T^N(F), T^N(F')) \cong s^*(\mathrm{Iso}(F, F'))\ .$$

But by A2.5.16 these are precisely what we need in order to construct a morphism $1_N \to \mathrm{Iso}(F, F')$. □

We remark that, in the above proof, the indexed category $\mathbb{E}$ could be replaced by any locally small $\mathcal{E}$-indexed category (that is, any locally internal category over $\mathcal{E}$ in the sense of B2.2.1).

Suggestions for further reading: Coste et al. [257], Freyd [371], Johnstone & Wraith [549], Loiseau [730], Rosebrugh [1034], Stout [1127].

D5.2 Finite cardinals

There are numerous possible notions of 'finiteness' for objects of a topos. Even in classical logic, it is well known that there are many different definitions of 'finite set', and that (at least some part of) the axiom of choice is required to prove all of them equivalent; in the non-classical internal logic of a topos, the range of possibilities is even wider, and we shall not in this work be able to explore more than a small fraction of those which have been considered. However, there are two particular notions which seem to be by far the most important for applications, and which we shall consider in the present section and Section D5.4 respectively.

We devote the present section to studying finite cardinals. The definition of finite cardinals was given in Section A2.5: a finite cardinal in a topos $\mathcal{E}$ (with a natural number object) is an object which occurs as a pullback, along a morphism $p\colon 1 \to N$, of the particular object $C = (s\pi_2\colon N_1 \to N)$ of $\mathcal{E}/N$, the *generic finite cardinal.* As in section A2.5, we shall write $[p]$ for the cardinal corresponding to a morphism $p\colon 1 \to N$. The key feature of cardinal-finiteness, which makes it such a good notion to work with, is precisely the fact that the finite cardinals in $\mathcal{E}$ are 'parametrized' by the generic finite cardinal, which in a sense contains all of them as particular cases; thus we can prove general results about finite cardinals by simply proving that they hold for the generic cardinal, and then observing that they are stable under pullback in an appropriate sense. We shall see many examples of this style of argument in what follows.

We saw in A2.5.14 that we have $[o] \cong 0$ and $[sp] \cong 1 \amalg [p]$ for any p. This gives rise to an important 'induction principle' for proving facts about finite cardinals, which we shall use repeatedly:

Lemma 5.2.1 *Let* P *be a property of objects which is expressible in the internal language of a topos, and suppose* (i) *that the initial object* 0 *satisfies* P, *and* (ii) *that whenever an object* A *satisfies* P, *so does* $1 \amalg A$. *Then every finite cardinal satisfies* P.

Proof Since P is expressible in the internal language, we can construct a subobject $N' \rightarrowtail N$ which is the truth-value of the assertion that the generic finite cardinal satisfies P in $\mathcal{E}/N$. The two conditions on P imply that N' is a (o, s)-closed subobject of N; so it is the whole of N. □

Condition (ii) of 5.2.1 may of course be split into two: that 1 satisfies P, and that the class of objects satisfying P is closed under binary coproducts. In this connection, we note

Lemma 5.2.2 *For natural numbers* p *and* q *in a topos* $\mathcal{E}$, *we have isomorphisms* $[p+q] \cong [p] \amalg [q]$, $[p \times q] \cong [p] \times [q]$ *and* $[p \leftarrow q] \cong [p]^{[q]}$. *In particular, the class of finite cardinals in* $\mathcal{E}$ *is closed under finite coproducts, finite products and exponentials.*

Proof Since the pullback functor q^* preserves coproducts, products and exponentials, it suffices to establish the three isomorphisms for the pair of natural numbers $(N^*p, 1_N)$ in $\mathcal{E}/N$. But we can do this using 5.1.9, by verifying that the objects on each side of the isomorphism satisfy the same recursion data. For example, for addition we have $[p+o] = [p] \cong [p] \amalg [o]$, $[p+sq] = [s(p+q)] \cong [p+q] \amalg 1$ and $[p] \amalg [sq] \cong [p] \amalg ([q] \amalg 1) \cong ([p] \amalg [q]) \amalg 1$, so that both sides are solutions of the data $o^*F \cong [p]$, $s^*F \cong F \amalg 1$. Similarly, the two sides of the second isomorphism satisfy $o^*F \cong 0$ and $s^*F \cong F \amalg [p]$; and those of the third isomorphism satisfy $o^*F \cong 1$ and $s^*F \cong F \times [p]$. The second assertion of the lemma is immediate from the first, plus the facts that $0 \cong [o]$ and $1 \cong [so]$ are finite cardinals. □

We write $\mathcal{E}_{fc}$ for the full subcategory of $\mathcal{E}$ whose objects are finite cardinals in $\mathcal{E}$. In addition to finite products and coproducts, it is closed under equalizers and coequalizers: to prove this, we require a couple more lemmas. Note first that cardinals are decidable, since they are subobjects of N; hence the equalizer of any parallel pair $[p] \rightrightarrows [q]$ is a complemented subobject of $[p]$.

Lemma 5.2.3 *A complemented subobject of a finite cardinal is a finite cardinal.*

Proof We note first that a complemented subobject of 1 is a cardinal: if $U \rightarrowtail 1$ has a complement $V \rightarrowtail 1$, we may identify it with $[p]$ where p is the unique morphism making the diagram

$$\begin{tikzcd} U \arrow[r, tail] \arrow[d] & 1 \arrow[d, "p"] & V \arrow[l, tail] \arrow[d] \\ 1 \arrow[r, "so"] & N & 1 \arrow[l, "o"'] \end{tikzcd}$$

commute.

Now consider the generic case. We have an object $D = (N^*2)^C$ in $\mathcal{E}/N$ indexing the complemented subobjects of the generic cardinal C; we have to construct a morphism $D \to N^*N$ whose cardinal is isomorphic to the generic complemented subobject of C. But constructing such a morphism is equivalent to constructing an element of the exponential $(N^*N)^D$ in $\mathcal{E}/N$; and we may do this by A2.5.16 using the morphism

$$1 \xrightarrow{o} N \cong N^1 \cong o^*((N^*N)^{(N^*2)^C})$$

and the morphism $(N^*N)^D \to s^*((N^*N)^D) \cong (N^*N)^{s^*D}$ whose exponential transpose is

$$(N^*N)^{(N^*2)^C} \times (N^*2)^C \times N^*2 \xrightarrow{\mathrm{ev} \times N^*u} N^*N \times N^*N \xrightarrow{+} N^*N$$

where $u\colon 2 \to N$ is the 'cardinality' of the generic complemented subobject of 1, constructed as above. It is then a straightforward application of 5.1.9 to verify that the cardinal of this natural number is isomorphic to the generic complemented subobject of C. □

In the opposite direction, we have

Lemma 5.2.4 *The image of any morphism $f\colon [p] \to [q]$ between finite cardinals is a complemented subobject of $[q]$.*

Proof Fix the cardinal $[q]$. The assertion, for an arbitrary object A, that the 'image' map $[q]^A \to \Omega^{[q]}$ factors through the subobject $2^{[q]} \rightarrowtail \Omega^{[q]}$ is expressible by a sentence in the internal language of $\mathcal{E}$. This sentence is trivially valid when $A = 0$, and its validity for $A = 1$ is just the assertion that $[q]$ is decidable. And if it holds for two objects A and B, then it holds for their coproduct, since a finite union of complemented subobjects is complemented. Hence by 5.2.1 it holds for all finite cardinals $[p]$. □

Lemma 5.2.5 *The coequalizer of a parallel pair of morphisms between finite cardinals is a finite cardinal.*

Proof As in 5.2.3, we consider first the case of a pair $f, g\colon 1 \rightrightarrows [q]$. By the last two lemmas, we have a coproduct decomposition

$$1 \xrightarrow{\ f\ } [q] \longleftarrow [r]\ ;$$

pulling this back along g, we get a coproduct decomposition $1 \cong U \amalg V$. Now the morphisms $1 \amalg g'\colon 1 \cong U \amalg V \to U \amalg [r]$ (where g' is the restriction of g) and $\nu_2\colon [r] \to U \amalg [r]$ induce a morphism $[q] \to U \amalg [r]$, which is readily seen to be a coequalizer of f and g; but U is a complemented subobject of 1 and hence a cardinal by 5.2.3, and so $U \amalg [r]$ is a cardinal by 5.2.2.

Now consider the general case: fix a cardinal $[p]$ in $\mathcal{E}$, and consider the assertion, about a variable object A, which says 'for each pair of morphisms $A \rightrightarrows [p]$, there exists $q\colon 1 \to N$ and a morphism $[p] \twoheadrightarrow [q]$ such that $A \rightrightarrows [p] \twoheadrightarrow [q]$ is a coequalizer'. This assertion is expressible in the internal language of $\mathcal{E}$, and it is clearly true for $A = 0$ since two morphisms $0 \rightrightarrows [p]$ must be equal. Suppose it holds for A, and consider a pair of morphisms $1 \amalg A \rightrightarrows [p]$. Then the coequalizer of $A \rightrightarrows [p]$ can be written in the form $[p] \twoheadrightarrow [q]$, and we have to form the coequalizer of the two composites $1 \rightrightarrows [p] \twoheadrightarrow [q]$. But by the first part of the proof this coequalizer is a finite cardinal; so $1 \amalg A$ satisfies the assertion above. Hence by 5.2.1 the assertion holds for all finite cardinals. □

Corollary 5.2.6 *A quotient of a finite cardinal is isomorphic to a finite cardinal iff it is decidable.*

Proof One direction is immediate from the fact that finite cardinals are decidable. Conversely, given an epimorphism $[p] \twoheadrightarrow A$ where A is decidable, its

kernel-pair is a complemented subobject of $[p] \times [p] \cong [p^2]$, and hence isomorphic to a finite cardinal by 5.2.3; but any epimorphism in a topos is the coequalizer of its kernel-pair, so the result follows from 5.2.5. □

Putting together the above results, we have

Theorem 5.2.7 *For any topos $\mathcal{E}$ with a natural number object, the full subcategory $\mathcal{E}_{fc}$ of finite cardinals in $\mathcal{E}$ is a Boolean topos. The inclusion functor $\mathcal{E}_{fc} \to \mathcal{E}$ preserves finite limits, finite colimits and exponentials; it is logical iff $\mathcal{E}$ is Boolean.*

Proof The existence of finite limits, finite colimits and exponentials, and their preservation by the inclusion functor, follow from 5.2.2, 5.2.3 and 5.2.5. By 5.2.3 and 5.2.4, the subobjects of a cardinal $[p]$ in $\mathcal{E}_{fc}$ are precisely its complemented subobjects in $\mathcal{E}$; so the cardinal $[sso] \cong 2$ will serve as a subobject classifier in $\mathcal{E}_{fc}$, and it is preserved by the inclusion functor iff $\mathcal{E}$ is Boolean. □

Examples 5.2.8 (*a*) Let $\mathcal{C}$ be a small category. Since inverse image functors preserve natural number objects by A2.5.6, the natural number object in $[\mathcal{C}, \mathbf{Set}]$ is simply the constant functor $\mathcal{C} \to \mathbf{Set}$ with value $\mathbb{N}$. Hence a natural number in $[\mathcal{C}, \mathbf{Set}]$ is a natural-number-valued function on ob $\mathcal{C}$ which is constant on each connected component of $\mathcal{C}$; and a finite cardinal is a finite-set-valued functor which is constant on each connected component. Thus $[\mathcal{C}, \mathbf{Set}]_{fc} \simeq (\mathbf{Set}_f)^I$, where I is the set of connected components of $\mathcal{C}$. (In particular, if G is a group, then $[G, \mathbf{Set}]_{fc}$ is equivalent to $\mathbf{Set}_f$, rather than (as one might perhaps have hoped) to the topos $[G, \mathbf{Set}_f]$ of A2.1.4; for the latter, one needs to take the topos $[G, \mathbf{Set}]_{dKf}$ of 5.4.18 below.)

(*b*) Similarly, if X is a topological space, the natural number object in $\mathbf{Sh}(X)$ is the constant sheaf $\mathbb{N}$, i.e. the sheaf of sections of the local homeomorphism $(\pi_1 \colon X \times \mathbb{N} \to X)$ (where $\mathbb{N}$ has the discrete topology); so a natural number in $\mathbf{Sh}(X)$ is a continuous (= locally constant) function $X \to \mathbb{N}$. If the space of components of X is discrete (for example if X is locally connected), then we may deduce that $\mathbf{Sh}(X)_{fc} \simeq (\mathbf{Set}_f)^I$, where I is the set of components of X.

Next, we note some further consequences of the induction principle 5.2.1.

Lemma 5.2.9

(i) *Finite cardinals are (internally) Dedekind-finite; that is, the canonical monomorphism* $\mathrm{Iso}([p],[p]) \rightarrowtail \mathrm{Mono}([p],[p])$ *is an isomorphism for all* p.

(ii) *Finite cardinals are internally projective in the sense of* 4.5.1(*b*); *that is, the functor* $(-)^{[p]}$ *preserves epimorphisms.*

(iii) *Finite cardinals are 'decidably well-ordered'; that is, if we order* $[p]$ *as a subobject of* N, *it satisfies the sentence which says that every complemented inhabited subobject of* $[p]$ *has a least member.*

Proof (i) The assertion that an object A is internally Dedekind-finite is expressible by a sentence in the internal language, and it is clearly true for 0 since $\mathrm{Iso}(0,0) \cong \mathrm{Mono}(0,0) \cong 1$. Suppose it is true for A, and suppose we have a monomorphism $m\colon 1 \amalg A \rightarrowtail 1 \amalg A$ (in some slice category $\mathcal{E}/B$ – but, to simplify the notation, we shall assume $B = 1$). Form the pullbacks

$$\begin{tikzcd} U_1 \amalg A_1 \arrow[r, tail] \arrow[d, tail, "m_1"] & 1 \amalg A \arrow[d, tail, "m"] & U_2 \amalg A_2 \arrow[l, tail] \arrow[d, tail, "m_2"] \\ 1 \arrow[r, tail, "\nu_1"] & 1 \amalg A & A \arrow[l, tail, "\nu_2"'] \end{tikzcd} .$$

Since m_1 is monic, we see that A_1 is subterminal, and indeed contained in the complement U_2 of U_1 in $\mathrm{Sub}(1)$; thus we have a monomorphism

$$A \cong A_1 \amalg A_2 \rightarrowtail U_2 \amalg A_2 \overset{m_2}{\rightarrowtail} A$$

which must be an isomorphism by hypothesis. It now follows easily that m_2 and $A_1 \rightarrowtail U_2$ are both isomorphisms, whence $m = m_1 \amalg m_2$ is an isomorphism.

(ii) The assertion that A is internally projective is not expressible by a sentence in the internal language, since it involves quantification over arbitrary epimorphisms in $\mathcal{E}$. However, the assertion, for a given epimorphism $e\colon B \twoheadrightarrow C$, that e^A is epic is so expressible. It is trivially satisfied when $A = 0$ or $A = 1$; and if it holds for A_1 and A_2 then it holds for $A_1 \amalg A_2$, since $(-)^{A_1 \amalg A_2} \cong (-)^{A_1} \times (-)^{A_2}$. So it holds for all finite cardinals, and for all epimorphisms in $\mathcal{E}$.

(iii) Here we are dealing with a property of ordered objects, rather than of objects pure and simple; but the induction principle works in the same way. Clearly 0 is (decidably) well-ordered. Suppose $(A, \leq)$ is decidably well-ordered, and consider a well-supported complemented subobject $B \rightarrowtail 1 \amalg A$, where (for the moment) we regard the extra element 1 as preceding all of A. We can decompose B as a coproduct $U \amalg A'$, where U is a complemented subterminal object; the fact that B is well supported ensures that $\sigma A'$ contains the complement V of U, and so V^*A' is a well-supported subobject of V^*A in $\mathcal{E}/V$. Hence we can find a morphism $a'\colon V \to A'$ which is its least element; then the unique morphism a making

$$\begin{tikzcd} U \arrow[r] \arrow[d] & 1 \arrow[d, "a"] & V \arrow[l] \arrow[d, "a'"] \\ U \arrow[r] & B & A' \arrow[l] \end{tikzcd}$$

commute is easily seen to be the least element of B.

Strictly speaking, we have proved that a finite cardinal is decidably well-ordered in the *opposite* of its usual ordering; if we identify [p] with the subobject

$\{x \colon N \mid sx \le p\}$, then the extra element of $[sp] \cong 1 \amalg [p]$ comes after the elements of $[p]$, not before them. However, any finite cardinal is order-isomorphic to its opposite; for 0 this is trivial, and for a successor cardinal $[sp]$ the isomorphism is given by the map $p \dot{-} (-)$. So $[p]$ is also decidably well-ordered in its usual ordering. □

We shall prove a stronger version of 5.2.9(ii) in the next section (see 5.3.10(i)).

Corollary 5.2.10 *In any topos $\mathcal{E}$ with a natural number object,*

(i) *the subcategory $\mathcal{E}_{fc}$ satisfies* (AC);

(ii) *the natural number object is decidably well-ordered.*

Proof (i) Lemma 5.2.9(ii) implies that $\mathcal{E}_{fc}$ satisfies (IC), since the inclusion $\mathcal{E}_{fc} \to \mathcal{E}$ preserves exponentials. However, as an object of $\mathcal{E}_{fc}$ the terminal object $1 \cong [so]$ is not just internally but externally projective; for if a cardinal $[p]$ is well-supported, then it has a (least) element by 5.2.9(iii). So the result follows from 4.5.4.

(ii) Let $A \rightarrowtail N$ be a well-supported complemented subobject. Since A is well-supported, we may assume (after pulling back along some $B \twoheadrightarrow 1$) that it contains some natural number p; so $A \cap [sp]$ is well-supported, and hence has a least element, which is clearly the least element of A. □

We may make the categories $(\mathcal{E}/I)_{fc}$, $I \in \text{ob } \mathcal{E}$, into an $\mathcal{E}$-indexed category $\mathbb{E}_f$, since pullback functors preserve finite cardinals. As an indexed category, $\mathbb{E}_f$ is 'small'; that is, it is equivalent to the standard indexing of an internal category, namely the internal full subcategory of the canonical indexing of $\mathcal{E}$ over itself generated by the generic finite cardinal (cf. B2.3.5). This internal category has already been exploited in B3.2.9; the following lemma is of use there, and also in Section D5.3.

Lemma 5.2.11 *Inverse image functors preserve exponentiation to the power of a finite cardinal; that is, if $f \colon \mathcal{F} \to \mathcal{E}$ is a geometric morphism, p a natural number in $\mathcal{E}$ and A an object of $\mathcal{E}$, then the canonical comparison map $f^*(A^{[p]}) \to (f^*A)^{[f^*p]}$ is an isomorphism.*

Proof It suffices to prove that $(f/N)^*$ preserves the generic such exponential $(N^*A)^C$. We can do this either by recalling that $\Sigma_N((N^*A)^C)$ is the list object LA (A2.5.17) and that list objects are preserved by cartesian functors having right adjoints, or alternatively using 5.1.9 and the fact that the two objects of $\mathcal{F}/N$ which we want to prove isomorphic satisfy the same recursion data. □

In dealing with finite cardinals in slice categories of $\mathcal{E}$, the following result is often useful.

Proposition 5.2.12 *Let $f \colon A \to [p]$ be a morphism of $\mathcal{E}$ whose codomain is a finite cardinal. Then f is (isomorphic to) a finite cardinal in $\mathcal{E}/[p]$ iff A is a finite cardinal in $\mathcal{E}$.*

Proof First suppose A is a finite cardinal $[q]$, say. Then we have a pullback diagram

$$\begin{tikzcd} {[q]} \arrow[r, "f"] \arrow[d, tail, "{(1,f)}"] & {[p]} \arrow[d, tail, "\Delta"] \\ {[q] \times [p]} \arrow[r, "f \times 1"] & {[p] \times [p]} \end{tikzcd}$$

in $\mathcal{E}$, which we may reinterpret as a pullback

$$\begin{tikzcd} f \arrow[r] \arrow[d, tail] & 1_{[p]} \arrow[d, tail] \\ {[p]^*[q]} \arrow[r, "{[p]^*f}"] & {[p]^*[p]} \end{tikzcd}$$

in $\mathcal{E}/[p]$. Since $(\mathcal{E}/[p])_{fc}$ is closed under pullbacks by 5.2.7, it follows that f is a finite cardinal.

Conversely, consider the assertion about an object B which says that, for every natural number $r \colon B \to N$ in $\mathcal{E}/B$, there exists $q \colon 1 \to N$ such that $\Sigma_B[r] \cong [q]$. This assertion is expressible in the internal language, and it is clearly true for $B = 0$ and $B = 1$. If it holds for B_1 and B_2, then for any $r = (r_1, r_2) \colon B_1 \amalg B_2 \to N$ it is easy to see that $\Sigma_B[r] \cong \Sigma_{B_1}[r_1] \amalg \Sigma_{B_2}[r_2]$, so it is a finite cardinal by 5.2.2. So the assertion holds for all finite cardinals, by 5.2.1. □

The next two results 'internalize' the idea that arbitrary coproducts may be constructed from finite coproducts and filtered colimits.

Lemma 5.2.13 *Let $\mathcal{S}$ be a topos with a natural number object, and let $\mathbb{C}$ and $\mathbb{D}$ be $\mathcal{S}$-indexed categories with $\mathcal{S}$-indexed coproducts, such that $\mathbb{D}$ is locally small and both $\mathbb{C}$ and $\mathbb{D}$ are stacks for the coherent coverage on $\mathcal{S}$. If $F \colon \mathbb{C} \to \mathbb{D}$ is an $\mathcal{S}$-indexed functor preserving (fibrewise) finite coproducts, then F also preserves coproducts indexed by finite cardinals in $\mathcal{S}$, i.e. the diagram*

$$\begin{tikzcd} \mathcal{C}^I \arrow[r, "F^I"] \arrow[d, "\Sigma_x"] & \mathcal{D}^I \arrow[d, "\Sigma_x"] \\ \mathcal{C}^J \arrow[r, "F^J"] & \mathcal{D}^J \end{tikzcd}$$

commutes up to isomorphism whenever $x \colon I \to J$ is a finite cardinal in $\mathcal{S}/J$.

Proof Since the Beck–Chevalley conditions hold for coproducts in both $\mathbb{C}$ and $\mathbb{D}$, it suffices to consider the case when x is the generic finite cardinal $C\colon N_1 \to N$. (Actually, we need to consider the generic cardinal in $\mathcal{S}/J$ for an arbitrary object J; but since the hypotheses are 'stable under slicing' we may as well take $J = 1$.) In this case, given an object A of $\mathcal{C}^{N_1}$, we have to construct an isomorphism $\Sigma_C F^{N_1}(A) \to F^N \Sigma_C(A)$ in $\mathcal{D}^N$; more precisely, to show that the canonical morphism between these two, regarded as a morphism $1 \to \mathbb{D}(\Sigma_C F^{N_1}(A), F^N \Sigma_C(A))$ in $\mathcal{S}/N$, factors through the object of isomorphisms $\mathrm{Iso}(\Sigma_C F^{N_1}(A), F^N \Sigma_C(A))$. To do this, we use 5.1.9 (or rather the extension of 5.1.9 mentioned in the remark which follows it). Applying o^* to this object of isomorphisms yields $\mathrm{Iso}(0_{\mathbb{D}}, F^1(0_{\mathbb{C}}))$ where $0_{\mathbb{C}}$ and $0_{\mathbb{D}}$ denote the initial objects of $\mathcal{C}^1$ and $\mathcal{D}^1$ respectively, since $\mathcal{C}^0 \simeq \mathcal{D}^0 \simeq \mathbf{1}$ by the descent condition, and this object is isomorphic to 1 since F^1 preserves the initial object. A similar argument using the fact that F^N preserves binary coproducts yields the required morphism

$$\mathrm{Iso}(\Sigma_C F^{N_1}(A), F^N \Sigma_C(A)) \longrightarrow s^*(\mathrm{Iso}(\Sigma_C F^{N_1}(A), F^N \Sigma_C(A)))\,.$$

□

Proposition 5.2.14 *In addition to the hypotheses of* 5.2.13, *suppose that* $\mathbb{C}$ *and* $\mathbb{D}$ *have, and* F *preserves, arbitrary* $\mathcal{S}$*-indexed filtered colimits. Then* F *preserves arbitrary* $\mathcal{S}$*-indexed coproducts.*

Proof The hypotheses on filtered colimits are to be taken, as in Sections C3.2 and C3.4, to refer to filtered internal categories in arbitrary slice categories $\mathcal{S}/J$ of $\mathcal{S}$; equivalently, by B2.6.6, to arbitrary weakly filtered categories in $\mathcal{S}$. In other words, the statement that F preserves filtered colimits means that, for a weakly filtered internal category $\mathbb{I}$ with $\pi_0\mathbb{I} \cong J$, the diagram

$$\begin{array}{ccc}
[\mathbb{I}, \mathcal{C}] & \xrightarrow{[\mathbb{I}, F]} & [\mathbb{I}, \mathcal{D}] \\
\Big\downarrow{\scriptstyle \varinjlim_{\mathbb{I}}} & & \Big\downarrow{\scriptstyle \varinjlim_{\mathbb{I}}} \\
\mathcal{C}^J & \xrightarrow{F^J} & \mathcal{D}^J
\end{array}$$

commutes up to canonical isomorphism. With this proviso, the hypotheses of the proposition are stable under slicing, so it suffices to show that for any object J of $\mathcal{S}$ the diagram

$$\begin{array}{ccc}
\mathcal{C}^J & \xrightarrow{F^J} & \mathcal{D}^J \\
\Big\downarrow{\scriptstyle \Sigma_J} & & \Big\downarrow{\scriptstyle \Sigma_J} \\
\mathcal{C}^1 & \xrightarrow{F^1} & \mathcal{D}^1
\end{array}$$

commutes.

The key idea is that we can express J as a filtered colimit of finite cardinals. In fact we have already done this in the proof of B3.2.9; let us recall that, in that example, we associated with an object J of $\mathcal{S}$ an $\mathbb{S}_f$-torsor $\mathbb{J}$ given by equipping the list object $L(J) = \Sigma_N((N^*J)^C)$ with an appropriate right action of $\mathbb{S}_f$. The assertion that $\mathbb{J}$ corresponds to a geometric morphism $\mathcal{S} \to [\mathbb{S}_f, \mathcal{S}]$ whose inverse image sends the generic object to J becomes, in this context, that the colimit in $\mathbb{S}$ of the diagram $\mathbb{J} \to \mathbb{S}_f \to \mathbb{S}$ (where the first factor is the discrete fibration, and the second is the inclusion functor) is J. Hence we may compute the J-indexed coproduct of an object of $\mathcal{C}^J$ by first computing the $L(J)$-indexed family of finite coproducts corresponding to this diagram, and then forming their filtered colimit. But F commutes up to isomorphism with both these operations (the first by 5.2.13, and the second by assumption); so it commutes with Σ_J, as required. □

Corollary 5.2.15 *Suppose $\mathbb{C}$ and $\mathbb{D}$ are cocomplete $\mathcal{S}$-indexed categories (where $\mathcal{S}$ is a topos with a natural number object), which are stacks for the coherent coverage on $\mathcal{S}$, and suppose that $\mathbb{D}$ is locally small. Then an $\mathcal{S}$-indexed functor $F: \mathcal{C} \to \mathcal{D}$ is $\mathcal{S}$-cocontinuous iff it preserves finite (fibrewise) colimits and filtered $\mathcal{S}$-indexed colimits.*

Proof This is immediate from 5.2.14 and the definition of cocontinuity (B1.4.13) for indexed functors. □

Suggestion for further reading: Johnstone & Wraith [549].

D5.3 Finitary algebraic theories

In this section we explore the relationship between finitary algebraic theories (in the sense of 1.1.7(a)) in toposes and monoids in the object classifier (cf. 3.2.1). We begin by recalling some facts about reflexive coequalizers (cf. A1.2.11).

Lemma 5.3.1 *Let $\mathcal{E}$ be a cartesian closed category, and $\mathbb{T}$ a finitary algebraic theory. Then the forgetful functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \to \mathcal{E}$ creates coequalizers of reflexive pairs. If in addition $\mathcal{E}$ has coequalizers of reflexive pairs, then the forgetful functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \to \mathcal{E}$ is monadic iff it has a left adjoint.*

Proof If $h, k: A \rightrightarrows B$ is a reflexive pair of $\mathbb{T}$-model homomorphisms and we are given a coequalizer $q: B \to C$ for (the underlying morphisms of) f and g in $\mathcal{E}$, then it follows from A1.2.12 (plus induction) that $A^n \rightrightarrows B^n \to C^n$ is a coequalizer for each natural number n. Hence, for each n-ary function symbol f of $\mathbb{T}$, the morphisms f_A and f_B induce a unique morphism $f_C: C^n \to C$ such that q is a homomorphism. Also, since q^n is epic for each n, C satisfies any equation in these operations which is satisfied by B; in particular, it is a $\mathbb{T}$-model. It is now straightforward to verify that, when we equip C with this $\mathbb{T}$-model structure, $q: B \to C$ becomes a coequalizer for h and k in $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$.

The second assertion is immediate from the first and A1.1.2, since the forgetful functor is clearly conservative. □

We shall also require the following properties of categories of $\mathbb{T}$-models:

Lemma 5.3.2 *Let $\mathbb{T}$ be a finitary algebraic theory.*

(i) *If $\mathcal{E}$ is regular (resp. effective regular), then $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$ is regular (resp. effective regular).*

(ii) *If $\mathcal{E}$ is locally cartesian closed, then the $\mathcal{E}$-indexed category $\mathbb{T}$-$\mathbb{M}\mathrm{od}(\mathcal{E})$ defined by $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})^I = \mathbb{T}$-$\mathbf{Mod}(\mathcal{E}/I)$ is $\mathcal{E}$-complete.*

(iii) *If $\mathcal{E}$ is locally cartesian closed and cocartesian, then $\mathbb{T}$-$\mathbb{M}\mathrm{od}(\mathcal{E})$ is $\mathcal{E}$-cocomplete provided the forgetful functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}/I) \to \mathcal{E}/I$ has a left adjoint for every I.*

Proof (i) Since image factorizations in $\mathcal{E}$ are stable under finite products, it is clear that the image in $\mathcal{E}$ of any homomorphism $h\colon A \to B$ of $\mathbb{T}$-models in $\mathcal{E}$ can be given a unique $\mathbb{T}$-model structure which makes both halves of the factorization of h into homomorphisms. It follows that the covers in $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$ are exactly the homomorphisms whose underlying $\mathcal{E}$-morphisms are covers; hence also $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$ has images, and they are stable under pullback. For effectiveness, we similarly verify that if $R \rightrightarrows A$ is an equivalence relation in $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$, then the coequalizer Q of $R \rightrightarrows A$ in $\mathcal{E}$ can be given a unique $\mathbb{T}$-model structure which makes the cover $A \twoheadrightarrow Q$ into a homomorphism; and $R \rightrightarrows A$ is the kernel-pair of $A \twoheadrightarrow Q$ in $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$, since it is so in $\mathcal{E}$ and the forgetful functor creates limits.

(ii) It is trivial that $\mathbb{T}$-$\mathbb{M}\mathrm{od}(\mathcal{E})$ has $\mathcal{E}$-indexed finite limits. For $\mathcal{E}$-indexed products, we have simply to observe that the pullback functors $x^*\colon \mathcal{E}/I \to \mathcal{E}/J$ induced by morphisms $x\colon J \to I$ in $\mathcal{E}$, and their right adjoints Π_x, both preserve finite limits and so lift to functors between the corresponding categories of $\mathbb{T}$-models.

(iii) The existence of finite colimits in the categories $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}/I)$, and their preservation by pullback functors, follows from 5.3.1 and a theorem of Linton [727]. The functors $\Sigma_x\colon \mathcal{E}/J \to \mathcal{E}/I$ do not lift directly to functors between the categories of $\mathbb{T}$-models, since they do not preserve finite products, but it follows from A1.1.3(i) that the 'lifted' left adjoints exist. □

If $\mathcal{E}$ satisfies the hypotheses of 5.3.2(i) and (ii) and A is a particular $\mathbb{T}$-model in $\mathcal{E}$, then we note that the $\mathcal{E}$-indexed category which to an object I associates the preorder of sub-$\mathbb{T}$-algebras of I^*A is $\mathcal{E}$-complete, since it is reflective in the $\mathcal{E}$-indexed slice category $\mathbb{T}$-$\mathbb{M}\mathrm{od}(\mathcal{E})/A$. In other words, we can form the internal intersection of an arbitrary $\mathcal{E}$-indexed family of subalgebras of a given algebra – a fact which we have already exploited in 5.1.1 (and in B4.2.10). Of course, it is not hard to prove directly that such intersections exist.

It is clear from 5.3.1 and 5.3.2 that we need to investigate the circumstances in which the forgetful functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \to \mathcal{E}$ has a left adjoint. There is one

case in which the construction of this left adjoint is particularly easy: namely, when $\mathbb{T}$ is the theory of monoids.

Proposition 5.3.3 *Let $\mathcal{E}$ be a topos with a natural number object. Then, for any object A of $\mathcal{E}$, the list object LA over A has the structure of a monoid; and in fact LA is the free monoid generated by A.*

Proof By 5.2.2, we have $[p+q] \cong [p] \amalg [q]$, from which it follows easily that $+^*(l\colon LA \to N) \cong \pi_1^*(l) \times \pi_2^*(l)$ in $\mathcal{E}/N \times N$, i.e. we have a pullback square

$$\begin{tikzcd} LA \times LA \arrow[r, "+_A"] \arrow[d, "l \times l"] & LA \arrow[d, "l"] \\ N \times N \arrow[r, "+"] & N \end{tikzcd} .$$

We take the top edge of this square to be the multiplication of the monoid LA; the unit is similarly defined by the pullback square

$$\begin{tikzcd} 1 \arrow[r, "o_A"] \arrow[d] & LA \arrow[d, "l"] \\ 1 \arrow[r, "o"] & N \end{tikzcd} .$$

The fact that $(LA, +_A, o_A)$ satisfies the axioms of the theory of monoids follows straightforwardly from the fact that $(N, +, o)$ does so. We note also that the construction of $+_A$ and o_A is 'natural' in A, so that L becomes a functor $\mathcal{E} \to$ **Mon**$(\mathcal{E})$.

To verify that L is left adjoint to the forgetful functor, we shall construct the unit and counit of the adjunction. The unit is the top edge of the pullback square

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[d] & LA \arrow[d, "l"] \\ 1 \arrow[r, "so"] & N \end{tikzcd} ;$$

again, this is clearly natural in A (and we note that $\eta_A = s_A(1 \times o_A)$). To construct the counit, let (M, m, e) be a monoid in $\mathcal{E}$; by the definition of list

objects, we have a unique morphism $\epsilon_M : LM \to M$ such that

$$\begin{tikzcd}
1 \arrow[r, "o_M"] \arrow[dr, "e"'] & LM \arrow[d, "\epsilon_M"] & M \times LM \arrow[l, "s_M"'] \arrow[d, "1 \times \epsilon_M"] \\
& M & M \times M \arrow[l, "m"]
\end{tikzcd}$$

commutes. Now we have $\epsilon_M \eta_M = \epsilon_M s_M (1 \times o_M) = m(1 \times \epsilon_M)(1 \times o_M) = m(1 \times e) = 1_M$, so we see that any morphism $f : A \to M$ from an arbitrary object A to the underlying object of a monoid M can be extended to a monoid homomorphism $\epsilon_M Lf : LA \to LM \to M$. But if we had another such extension g, the equalizer of g and $\epsilon_M Lf$ would be a submonoid of LA containing η_A, from which it is easy to verify that it would be a (o_A, s_A)-closed subobject and hence the whole of LA (cf. 5.1.5). So the extension is unique, and the adjunction is established. □

Remark 5.3.4 In fact the existence of a natural number object in $\mathcal{E}$ is necessary as well as sufficient for the existence of a free functor $\mathcal{E} \to \mathbf{Mon}(\mathcal{E})$. We can see this using 5.1.3: for if (M, m, e) is a monoid freely generated by 1 (with $g : 1 \to M$ as the unit of the adjunction), then it is easily verified that $1 \amalg M$ is also a monoid with unit ν_1 and multiplication

$$(1 \amalg M) \times (1 \amalg M) \cong 1 \amalg M \amalg M \amalg (M \times M) \xrightarrow{(\nu_1, \nu_2, \nu_2, \nu_2 sm)} 1 \amalg M,$$

where s is the composite

$$M \xrightarrow{1 \times g} M \times M \xrightarrow{m} M,$$

and that $1 \amalg M$ is freely generated by the element $\nu_2 e : 1 \to 1 \amalg M$. So we have an isomorphism $M \cong 1 \amalg M$, and hence by 5.1.3 $\mathcal{E}$ has a natural number object – which is of course M itself. (Alternatively, we could argue as in the proof of 5.4.10 below to show that a free monoid generated by A necessarily has the universal property of a list object over A.)

Variants of this argument can be given for a number of other familiar algebraic theories (for example, for groups); but there are also some nontrivial theories, such as that of semilattices (cf. 5.4.9 below), for which the free functor exists in any topos.

Theorem 5.3.5 *Let $\mathcal{E}$ be a topos with a natural number object, and let $\mathbb{T}$ be a finitely-presented finitary algebraic theory (i.e. one whose signature has finitely many function symbols, and which has finitely many axioms). Then the forgetful functor $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \to \mathcal{E}$ has a left adjoint.*

Proof We consider first the case when $\mathbb{T}$ is a free theory, i.e. one with no axioms. Let $F = \{f_1, f_2, \ldots, f_k\}$ be the finite set of function symbols of $\mathbb{T}$; as usual, we identify F with a finite copower of 1 in $\mathcal{E}$, and write $\alpha\colon F \to N$ for the morphism indexing each operation by its arity (regarded as a standard natural number in $\mathcal{E}$). We shall construct the free $\mathbb{T}$-model on an object A of $\mathcal{E}$ as a subobject of the list object $L = L(F \amalg A)$.

We note that L itself has a $\mathbb{T}$-model structure induced by its monoid structure: given an operation f_i (with $\alpha(f_i) = n_i$, say), we define $(f_i)_L\colon L^{n_i} \to L$ to be the interpretation of the term $\eta\nu_1(f_i) + x_1 + \cdots + x_{n_i}$, where η and $+$ are as defined in the proof of 5.3.3, and $x_1, \ldots, x_{n_i}$ are variables of sort L. Also, we have a morphism $\eta\nu_2\colon A \to L$.

Now let $N_\perp$ denote $1 \amalg N$, with the element $\nu_1\colon 1 \to 1 \amalg N$ denoted by $\perp$ (and elements of N identified with their images under ν_2). We define $p\colon L \to N_\perp$ to be the unique morphism making the diagram

$$\begin{tikzcd} 1 \arrow[r, "o"] \arrow[dr, "o"'] & L \arrow[d, "p"] & (F \amalg A) \times L \arrow[l, "s"'] \arrow[d, "1 \times p"] \\ & N_\perp & (F \amalg A) \times N_\perp \arrow[l, "t"'] \end{tikzcd}$$

commute, where t is the morphism defined by

$$t(x,y) = \begin{cases} \perp & \text{if } y = \perp \\ sy & \text{if } y \in N \text{ and } x \in A \\ s(y \dot{-} \alpha x) & \text{if } y \in N,\ x \in F \text{ and } \alpha x \leq y \\ \perp & \text{if } y \in N,\ x \in F \text{ and } \alpha x > y\,. \end{cases} \tag{5.1}$$

(Intuitively, p is the morphism which sends a string of elements of $F \amalg A$ to n if the string can be 'parsed' as a sequence of n 'well-formed words', and to $\perp$ if it cannot be so parsed.) It is now straightforward to verify that the pullback

$$\begin{tikzcd} M \arrow[r] \arrow[d, tail] & 1 \arrow[d, tail, "so"] \\ L \arrow[r, "p"] & N_\perp \end{tikzcd}$$

is a sub-$\mathbb{T}$-model of L, and that it contains $\eta\nu_2\colon A \rightarrowtail L$; also, the restrictions of the $(f_i)_L\colon L^{n_i} \to L$ to M^{n_i} are all monic, since if $p(\eta\nu_1(f_i) + x) = so$, then $p(x) = n_i$ and there is a unique way of decomposing x as a sum of n_i elements of M. And the images of these morphisms are disjoint from each other and from $\eta\nu_2$, since this is already true for L. It is also easy to see that these morphisms

are jointly epic, so in fact we have a coproduct decomposition $M \cong A \amalg M^{n_1} \amalg M^{n_2} \amalg \ldots \amalg M^{n_k}$.

It can be shown without much difficulty that M is generated as a $\mathbb{T}$-model by $\eta\nu_2 \colon A \rightarrowtail M$ (i.e. that it is the internal intersection of all sub-$\mathbb{T}$-models containing this subobject); but even without verifying this we could cut down to the submodel generated by A by the remarks after 5.3.2, and then proceed much as in the proof of 5.1.2(iii) $\Rightarrow$ (i). Given a $\mathbb{T}$-model B, together with a morphism $h \colon A \to B$, we consider the smallest sub-$\mathbb{T}$-model $(a, b) \colon R \rightarrowtail M \times B$ which contains the image of $(g, h) \colon A \to M \times B$, and verify that R is the graph of a homomorphism $M \to B$ by showing that $\{x \colon M \mid (\exists! y \colon R)(a(y) = x)\}$ defines a sub-$\mathbb{T}$-model of M containing A. Hence every such h extends to a homomorphism $M \to B$; but the equalizer of any two such homomorphisms would be a sub-$\mathbb{T}$-model of M containing A, and so the extension is unique.

We now prove the general case by induction on the number of axioms of $\mathbb{T}$. Suppose that $\mathbb{T}$ is obtained by adding a single axiom to a theory $\mathbb{S}$, and we already know that free $\mathbb{S}$-models exist. We shall show that $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E})$ is reflective in $\mathbb{S}$-$\mathbf{Mod}(\mathcal{E})$. Suppose the extra equation in $\mathbb{T}$ involves n free variables; then for each $\mathbb{S}$-model B we may construct a pair of morphisms $B^n \rightrightarrows B$ in $\mathcal{E}$ corresponding to the terms on either side of the equation, such that these morphisms are equal iff B is a $\mathbb{T}$-model. Let F denote the free $\mathbb{S}$-model functor, and form the coequalizer $q \colon B \to C$ of the induced pair $F(B^n) \rightrightarrows B$ in $\mathbb{S}$-$\mathbf{Mod}(\mathcal{E})$ (which we may do by 5.3.2(iii)). By the proof of 5.3.2(i), we know that q is epic in $\mathcal{E}$, and hence that $q^n \colon B^n \to C^n$ is epic; but since q is a homomorphism it commutes with the morphisms $B^n \rightrightarrows B$ and $C^n \rightrightarrows C$ which are the interpretations of the terms in the extra equation, and hence C must satisfy this equation, i.e. it is a $\mathbb{T}$-model. But any homomorphism $h \colon B \to D$ from B to a $\mathbb{T}$-model must have equal composites with $B^n \rightrightarrows B$ and hence with $F(B^n) \rightrightarrows B$; so it factors uniquely through q. In particular, if B is the free $\mathbb{S}$-model on some object A of $\mathcal{E}$, then C has the universal property of a free $\mathbb{T}$-model on A. □

We note that, in both halves of the proof of 5.3.5, the finite presentability assumption could be replaced by an '$\mathcal{E}$-presentability' assumption: in the first part, we could replace the finite set F by an arbitrary object of $\mathcal{E}$ equipped with an 'arity' map $F \to N$, and in the second we could replace the single equation by an 'I-indexed family of equations'. In particular, if $\mathcal{E}$ is cocomplete, we can construct free functors for arbitrary (**Set**-based) finitary algebraic theories; but in fact the assertion we have just made is more general than this, as we shall see below.

Remark 5.3.6 There is an alternative, less explicit, proof of 5.3.5, using the ideas of B4.2.10. Given a finitely-presented algebraic theory $\mathbb{T}$ and an object A of $\mathcal{E}$, we may consider the geometric theory $(A \downarrow \mathbb{T})$ over $\mathcal{E}$ (in the sense of B4.2.7) whose models consist of a $\mathbb{T}$-model M together with a morphism from A to the underlying object of M. We may construct a classifying topos $(p \colon \mathcal{E}[(A \downarrow \mathbb{T})] \to \mathcal{E})$ for this theory, containing a generic model $(f \colon p^*A \to M)$,

by the method of B4.2.9. Then, as in the proof of B4.2.10, we may show that the transpose $\overline{f}\colon A \to p_*M$ of f is a weakly initial object in the category of $(A \downarrow \mathbb{T})$-models in $\mathcal{E}$; and, by cutting down to the internal intersection of all sub-$\mathbb{T}$-models of p_*M which contain the image of $\overline{f}$, we may obtain an initial object of this category, that is a free $\mathbb{T}$-model generated by A.

Next, we note

Lemma 5.3.7 *Let $f\colon \mathcal{F} \to \mathcal{E}$ be a geometric morphism, and $\mathbb{T}$ an algebraic theory such that the forgetful functors $\mathbb{T}$-$\mathbf{Mod}(\mathcal{E}) \to \mathcal{E}$ and $\mathbb{T}$-$\mathbf{Mod}(\mathcal{F}) \to \mathcal{F}$ have left adjoints (for example, suppose that $\mathbb{T}$ is finitely-presented and that $\mathcal{E}$ (and hence $\mathcal{F}$) has a natural number object). Then the diagram*

$$\begin{tikzcd} \mathcal{E} \arrow[r, "f^*"] \arrow[d] & \mathcal{F} \arrow[d] \\ \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E}) \arrow[r, "f^*"] & \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{F}) \end{tikzcd}$$

commutes up to natural isomorphism, where the vertical arrows denote the free functors.

Proof The functors in the diagram all have right adjoints, and the diagram of right adjoints trivially commutes. So this is immediate from uniqueness of adjoints. □

Corollary 5.3.8 *Let $\mathcal{E}$ be a topos with a natural number object, and $\mathbb{T}$ a finitely-presented algebraic theory. For each $\mathcal{E}$-topos $(f\colon \mathcal{F} \to \mathcal{E})$, let $\mathbb{T}_{\mathcal{F}} = (T_{\mathcal{F}}, \eta_{\mathcal{F}}, \mu_{\mathcal{F}})$ be the monad on $\mathcal{F}$ induced by the free $\mathbb{T}$-algebra functor and its right adjoint. Then the monads $\mathbb{T}_{\mathcal{F}}$ form a $\mathfrak{Top}/\mathcal{E}$-indexed monad on the $\mathfrak{Top}/\mathcal{E}$-indexed category $(f\colon \mathcal{F} \to \mathcal{E}) \mapsto \mathcal{F}$.*

Proof This is immediate from 5.3.7 and the fact that inverse image functors also commute with the forgetful functors $\mathbb{T}$-$\mathbf{Mod}(\mathcal{F}) \to \mathcal{F}$. □

We have seen indexed monads of the type described in 5.3.8 before. In 3.2.1 we defined a monoidal structure on the object classifier $\mathbf{Set}[\mathbb{O}]$, and observed that monoids for this structure correspond to $\mathfrak{Top}/\mathbf{Set}$-indexed monads. The same argument works over any base topos $\mathcal{E}$ with a natural number object: if $\mathcal{E}[\mathbb{O}]$ denotes the object classifier over $\mathcal{E}$ (which exists by B3.2.9), then an indexed endofunctor of $\mathcal{E}$-toposes corresponds to an object $T = T_{\mathcal{E}[\mathbb{O}]}(G_{\mathbb{O}})$ of $\mathcal{E}[\mathbb{O}]$ (where $G_{\mathbb{O}}$ is the generic object), and an indexed monad structure on the functors $T_{\mathcal{F}}$ corresponds to a monoid structure on T. Accordingly, we now introduce the

following definition:

Definition 5.3.9 Let $\mathcal{E}$ be a topos with a natural number object. By an (*internal*) *finitary algebraic theory* in $\mathcal{E}$, we mean a monoid in the monoidal category $(\mathcal{E}[\mathbb{O}], \otimes, G_{\mathbb{O}})$ of 3.2.1. If $\mathbb{T} = (T, m, e)$ is such a monoid, we write $\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{F})$ (where $\mathcal{F}$ is an arbitrary topos defined over $\mathcal{E}$) for the category of algebras of the monad whose functor part is the $\mathcal{F}$-component $T_{\mathcal{F}}$ of the indexed endofunctor of $\mathcal{E}$-toposes corresponding to T.

We have seen that, in the case $\mathcal{E} = \mathbf{Set}$, the above definition includes the classical notion of finitary algebraic theory as a special case. (In fact, as we shall see below, it is equivalent to the latter.) We next establish the analogue of 5.3.1 for our new notion:

Lemma 5.3.10 *Let $\mathcal{E}$ be a topos with a natural number object.*

(i) *If $p\colon 1 \to N$ is a natural number in $\mathcal{E}$, then the functor $(-)^{[p]}\colon \mathcal{E} \to \mathcal{E}$ preserves coequalizers of reflexive pairs.*
(ii) *If T is an object of $\mathcal{E}[\mathbb{O}]$, then for any $\mathcal{E}$-topos $\mathcal{F}$ the functor $T_{\mathcal{F}}\colon \mathcal{F} \to \mathcal{F}$ preserves coequalizers of reflexive pairs.*
(iii) *If $\mathbb{T}$ is a finitary algebraic theory in $\mathcal{E}$ (in the sense of 5.3.9), then for any $\mathcal{E}$-topos $\mathcal{F}$ the forgetful functor $\mathbb{T}\text{-}\mathbf{Mod}(\mathcal{F}) \to \mathcal{F}$ creates coequalizers of reflexive pairs.*

Proof (i) This is proved by the same method as 5.2.9(ii): given a particular reflexive coequalizer diagram $A \rightrightarrows B \twoheadrightarrow C$ in $\mathcal{E}$, the assertion that an exponential functor $(-)^X$ preserves it is expressible by a sentence in the internal language of $\mathcal{E}$. This sentence is clearly satisfied when $X = 0$ or $X = 1$; and if it holds for X_1 and X_2 then A1.2.12 ensures that it holds for $X_1 \amalg X_2$. So it holds for all finite cardinals.

(ii) Since the hypotheses are stable under change of base from $\mathcal{E}$ to $\mathcal{F}$, it suffices to consider the case $\mathcal{E} = \mathcal{F}$. We recall from B3.2.9 that the object classifier $\mathcal{E}[\mathbb{O}]$ may be constructed as the internal diagram category $[\mathbb{E}_f, \mathcal{E}]$, and that the object of objects of $\mathbb{E}_f$ is N. Hence by B2.3.16 $\mathcal{E}[\mathbb{O}]$ is monadic over $\mathcal{E}/N$. In particular, any object T of $\mathcal{E}[\mathbb{O}]$ admits a free presentation, that is a (reflexive) coequalizer diagram

$$\mathbb{R}(r) \rightrightarrows \mathbb{R}(g) \longrightarrow T,$$

where $g\colon G \to N$ and $r\colon R \to N$ are objects of 'generators' and 'relations' in $\mathcal{E}/N$ and $\mathbb{R}$ is the left adjoint to the forgetful functor (cf. B2.5.4(c)). But, for any object A of $\mathcal{E}$, the functor $T \mapsto T_{\mathcal{E}}(A)$ is the inverse image $\overline{A}^*$ of the geometric morphism $\overline{A}\colon \mathcal{E} \to \mathcal{E}[\mathbb{O}]$, and hence preserves coequalizers; so we have

a coequalizer diagram

$$\mathbb{R}(r)_{\mathcal{E}}(A) \rightrightarrows \mathbb{R}(g)_{\mathcal{E}}(A) \longrightarrow T_{\mathcal{E}}(A)\,.$$

By the construction of B3.2.9, the $\mathbb{E}_{fc}$-torsor corresponding to A has $(N^*A)^C$ for its underlying object, where C is the generic cardinal in $\mathcal{E}/N$, and it follows easily that, for any $g\colon G \to N$, the functor $\mathbb{R}(g)_{\mathcal{E}}$ may be identified with the composite

$$\mathcal{E} \xrightarrow{\ G^*\ } \mathcal{E}/G \xrightarrow{(-)^{[g]}} \mathcal{E}/G \xrightarrow{\ \Sigma_G\ } \mathcal{E},$$

i.e. the partial product functor $P(-,g)$ (cf. A1.5.7). And this preserves reflexive coequalizers, by (i) and the fact that G^* and Σ_G have right adjoints; so the result for $T_{\mathcal{E}}$ follows by a straightforward diagram-chase.

(iii) is immediate from (ii) and the definition of $\mathbb{T}$-**Mod**$(\mathcal{F})$. □

Corollary 5.3.11 *Let $\mathcal{E}$, $\mathbb{T}$ and $\mathcal{F}$ be as in* 5.3.10(iii). *Then* $\mathbb{T}$-**Mod**$(\mathcal{F})$ *is effective regular and cocartesian.*

Proof The forgetful functor $\mathbb{T}$-**Mod**$(\mathcal{F}) \to \mathcal{F}$ creates limits (in particular, kernel-pairs); so 5.3.10(iii) enables us to lift image factorizations from $\mathcal{F}$ to $\mathbb{T}$-**Mod**$(\mathcal{F})$, and to check that they are stable under pullback and that equivalence relations in $\mathbb{T}$-**Mod**$(\mathcal{F})$ are kernel-pairs of their coequalizers. The existence of arbitrary finite colimits in $\mathbb{T}$-**Mod**$(\mathcal{F})$ follows from Linton's theorem, as in 5.3.2(iii). □

Remark 5.3.12 In the case $\mathcal{E} = \mathbf{Set}$, the functor $T_{\mathbf{Set}}\colon \mathbf{Set} \to \mathbf{Set}$ induced by an object T of $\mathbf{Set}[\mathbb{O}] = [\mathbf{Set}_f, \mathbf{Set}]$ may be alternatively described as the left Kan extension of T along the inclusion $\mathbf{Set}_f \to \mathbf{Set}$ (cf. the discussion after 3.2.1, or alternatively the proof of 5.3.10(ii)). Thus a $\otimes$-monoid in $\mathbf{Set}[\mathbb{O}]$ induces a monad on **Set** whose functor part is finitary; but this in turn can be presented as a (finitary) algebraic theory in the usual sense, taking the elements of $T(n)$ as n-ary function symbols for each natural number n, with appropriate equations. Combining this with 5.3.8, we see that internal finitary algebraic theories in **Set** are (up to equivalence) the same thing as algebraic theories in the classical sense. However, the advantage of Definition 5.3.9 is that it enables us to give a unified treatment of these theories and of others whose operations and equations 'vary within a topos': for example, if R is an internal ring in a topos $\mathcal{E}$, we have a notion of R-module (that is, of $f^*(R)$-module) in any $\mathcal{E}$-topos $f\colon \mathcal{F} \to \mathcal{E}$; provided $\mathcal{E}$ has a natural number object, it is easy to construct free R-modules, by an appropriate modification of the construction in 5.3.5, and hence to express R-modules as the models of an internal finitary algebraic theory in $\mathcal{E}$.

If T is an object of $\mathcal{E}[\mathbb{O}]$, then $(-) \otimes T$ and $T \otimes (-)$ both preserve reflexive coequalizers as functors $\mathcal{E}[\mathbb{O}] \to \mathcal{E}[\mathbb{O}]$: the former because it has a right adjoint by 3.2.1, and the latter by 5.3.10(ii). From this, it follows by A1.2.12 (and the argument in the proof of 5.3.1) that the forgetful functor $\mathbf{Mon}(\mathcal{E}[\mathbb{O}], \otimes, G_{\mathbb{O}}) \to \mathcal{E}[\mathbb{O}]$

creates reflexive coequalizers. Given that it also has a left adjoint (which may be constructed by an appropriate recursion – we refer to [549] for the details), it follows by A1.1.2 that it is monadic; and so also is the composite forgetful functor

$$\mathbf{Mon}(\mathcal{E}[\mathbb{O}], \otimes, G_{\mathbb{O}}) \longrightarrow \mathcal{E}[\mathbb{O}] \longrightarrow \mathcal{E}/N .$$

Thus we obtain the idea of a *free internal finitary algebraic theory* generated by an object $g: G \to N$ of $\mathcal{E}/N$ (that is, by a G-indexed family of 'finitary operations', together with a morphism specifying their arities), and the fact that any internal finitary algebraic theory has a 'free presentation' as a coequalizer of a pair of morphisms between such free theories. We shall not pursue this topic further: the reader is referred to [549] for a more detailed account.

Suggestions for further reading: Blass [124], Johnstone & Wraith [549], Lesaffre [719, 720], Rosebrugh [1032, 1033], Schumacher [1099].

D5.4 Kuratowski-finiteness

The notion of finite cardinal, studied in Section D5.2, provides us with a notion of 'finite object' in any topos with a natural number object, and as we have seen it enables us to develop a lot of 'finitary mathematics' inside such a topos. However, there is another notion of finiteness, which has the advantage that it can be defined in any topos (without assuming a natural number object), and which is more appropriate for many purposes – particularly in contexts involving ordered objects and lattices (cf. 5.4.26 below). This notion first appeared, in the context of set theory, in a paper by W. Sierpiński [1109]; but the first detailed investigation of it was made by C. Kuratowski [655], and it has become generally known as Kuratowski-finiteness (or K-finiteness for short).

We saw in 5.2.1 that cardinal-finiteness corresponds to a certain 'induction principle': if we can prove (in the internal logic of our topos) that some assertion holds for the particular finite cardinals $0 = [o]$ and $1 = [so]$, and that it holds for a binary coproduct (= disjoint union) $A \amalg B$ provided it holds for both A and B, then it holds for all finite cardinals. Kuratowski-finiteness corresponds in the same way to a very similar induction principle: the only formal difference is that we replace the disjoint unions $A \amalg B$ by arbitrary binary unions $A \cup B$ of subobjects. However, the crucial difference is that the Kuratowski induction principle is 'local'; that is, it takes place inside the poset of subobjects of some fixed object of the topos, rather than in the topos as a whole. This is the reason why the Kuratowskian definition works without a natural number object: we can reasonably expect that the collection of all finite subobjects of some given object A should be 'parametrized' by some object $K(A)$ of the topos (we shall shortly see how to construct it as a subobject of PA) – whereas the existence of an object F parametrizing the collection of all finite objects of the topos, for any reasonable definition of finiteness, will imply the existence of a natural number

object, since we shall have morphisms $1 \to F$ (corresponding to the object 0) and $F \rightarrowtail F$ (corresponding to the operation $A \mapsto A \amalg 1$) whose images are clearly disjoint. (In fact, even if our topos $\mathcal{E}$ has a natural number object, we shall not in general be able to construct an object parametrizing *all* the K-finite objects of $\mathcal{E}$ in this sense.)

The starting-point of the theory is thus the definition of the object $K(A)$ for an arbitrary object A of $\mathcal{E}$.

Definition 5.4.1

(*a*) Let A be an object of a topos $\mathcal{E}$. We define $K(A) \rightarrowtail PA$ to be the sub-join-semilattice of PA generated by the subobject $\{\}: A \rightarrowtail PA$: in other words, it is the internal intersection of all subobjects B of PA which contain $\{\}$ and $\ulcorner 0 \urcorner : 1 \to PA$ (the name of the subobject $0 \rightarrowtail A$), and are closed under the binary join map $\cup: PA \times PA \to PA$. (Note that, in this and the next section, we shall generally write $\cup$ rather than $\vee$ for the binary join map, in order to distinguish it from disjunction in the internal language of $\mathcal{E}$.)

(*b*) We say A is *Kuratowski-finite* (or simply *K-finite*) if the top element $\top_A : 1 \to PA$ of PA factors through $K(A) \rightarrowtail PA$.

We shall see before long that $K(A)$ is indeed the object of K-finite subobjects of A, in the sense that a morphism $1 \to PA$ factors through $K(A)$ iff (the domain of) the subobject $A' \rightarrowtail A$ which it names is a K-finite object. However, it is *not* true in general that every subobject of a K-finite object is K-finite; thus we do not assert that A is K-finite iff $K(A)$ is the whole of PA. (Again, we shall see a counterexample before long.) On the other hand, saying that A is K-finite *is* equivalent to saying that the internal poset $K(A)$ has a greatest element; for such an element must be an (internal) upper bound for the singletons of A, and hence it must be $\top_A$.

Remark 5.4.2· $K(A)$ could equivalently be defined as the smallest subobject of PA which contains $\ulcorner 0 \urcorner : 1 \to PA$ and is closed under the operation of 'adding singletons', in the sense that

$$A \times PA \xrightarrow{\{\} \times 1} PA \times PA \xrightarrow{\cup} PA$$

restricts to a morphism $A \times K(A) \to K(A)$ (which we shall denote by σ_A). For $K(A)$ clearly has these closure properties; conversely, if we write $K'(A)$ for the smallest subobject with these closure properties, then we may show that $K'(A)$ is a join-semilattice (and hence has the closure properties of $K(A)$), by showing that the subobject of $K'(A)$ defined by the term

$$\{w : K'(A) \mid (\forall w' : K'(A))((w \cup w') \in \ulcorner K'(A) \urcorner)\}$$

has the closure properties of $K'(A)$ and hence is the whole of $K'(A)$.

Lemma 5.4.3

(i) *The assignment $A \mapsto K(A)$ defines a functor $\mathcal{E} \to \mathcal{E}$ (in fact a subfunctor of the covariant power-object functor).*

(ii) *The functor K preserves monomorphisms and epimorphisms.*

Proof (i) If $f: A \to B$ is a morphism of $\mathcal{E}$, then the diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, tail, "\{\}"] & B \arrow[d, tail, "\{\}"] \\ PA \arrow[r, "\exists f"] & PB \end{tikzcd}$$

commutes, as we saw in A2.3.6(i), and $\exists f$ (being an internal left adjoint) is a join-semilattice homomorphism; so it restricts to a morphism $K(f): K(A) \to K(B)$. Thus K is a subfunctor of the covariant power-object functor.

(ii) The fact that K preserves monomorphisms is immediate from the fact that the covariant power-object functor does so (cf. A2.2.5). For epimorphisms, let $f: A \twoheadrightarrow B$ be an epimorphism. The image of $K(A)$ under $\exists f: PA \to PB$ is a sub-join-semilattice of PB, since $\exists f$ is a join-semilattice homomorphism; and it contains $\{\}: B \rightarrowtail PB$ by the commutativity of the diagram in the proof of (i) and the fact that f is epic. So it contains $K(B)$; but the reverse inclusion is simply the functoriality of K. □

We may now establish some of the basic closure properties of the class of K-finite objects.

Lemma 5.4.4

(i) *The objects 0 and 1 are K-finite.*

(ii) *Epimorphic images of K-finite objects are K-finite.*

(iii) *An object which is a union of two K-finite subobjects is K-finite.*

(iv) *A coproduct $A \amalg B$ is K-finite iff both A and B are K-finite.*

Proof (i) We have $P(0) \cong 1$, so its top and bottom elements coincide; $K(0)$ contains the bottom element by definition, so it contains the top element. For 1, we claim that $K(1)$ is the subobject $(\bot, \top): 1 \amalg 1 \rightarrowtail \Omega$; for it contains these two elements of $P(1) \cong \Omega$ (note that $\top$ is the singleton map $1 \to P(1)$), and they form a sub-join-semilattice. In particular, it contains the top element of Ω.

(ii) If $f: A \to B$ is epimorphic, then the composite $\exists f \circ \top_A: 1 \to PA \to PB$ equals $\top_B$; so this is immediate from 5.4.3(i).

(iii) Similarly, for a monomorphism $m: A \rightarrowtail B$, we have $\exists m \circ \top_A = \ulcorner m \urcorner$, so if B is the union of m and $m': A' \rightarrowtail B$ where A and A' are K-finite, then both $\ulcorner m \urcorner$ and $\ulcorner m' \urcorner$ factor through $K(B)$. But since $K(B)$ is a join-semilattice, so does the name of their union, i.e. $\top_B$.

(iv) One direction follows immediately from (iii). Conversely, suppose $A \amalg B$ is K-finite. It is easily seen that the square

$$\begin{tikzcd} A \amalg B \arrow[r] \arrow[d, "\{\}"] & A \amalg 1 \arrow[d, "(\{\}, \ulcorner 0 \urcorner)"] \\ P(A \amalg B) \arrow[r, "P\nu_1"] & PA \end{tikzcd}$$

commutes, since both ways round correspond to the subobject $(\nu_1, 1_A)\colon A \rightarrowtail (A \amalg B) \times A$. But $P\nu_1$ is a join-semilattice homomorphism since pullback preserves unions of subobjects; and since $K(A)$ is a semilattice it contains the element $\ulcorner 0 \urcorner$. Hence $P\nu_1$ maps $K(A \amalg B)$ into $K(A)$, and so $\top_A = P\nu_1 \circ \top_{A \amalg B}$ factors through $K(A)$. □

Corollary 5.4.5

(i) *$K(A)$ is 'the object of K-finite subobjects of A'; i.e., given a subobject $m\colon A' \rightarrowtail A$, A' is K-finite iff $\ulcorner m \urcorner\colon 1 \to PA$ factors through PA.*

(ii) *A subterminal object is K-finite iff it is complemented.*

(iii) *For any two objects A and B, we have $K(A \amalg B) \cong K(A) \times K(B)$.*

Proof (i) If A' is K-finite, then $\ulcorner m \urcorner = \exists m \circ \top_{A'}$ factors through $K(A)$ by the argument of 5.4.4(iii). For the converse, we observe that the assertion 'A is K-finite' is expressible by a sentence in the internal language of $\mathcal{E}$, so we may construct 'the object of K-finite subobjects of A' as a subobject $\hat{K}(A)$ of PA, by evaluating the truth-value of this sentence for the generic subobject of $PA^*(A)$ in $\mathcal{E}/PA$. Then parts (i) and (iii) of 5.4.4 ensure that $\hat{K}(A)$ is a sub-join-semilattice of PA and that it contains $\{\}\colon A \rightarrowtail PA$; so it contains $K(A)$.

(ii) follows easily from (i) and the fact, which we observed in the proof of 5.4.4(i), that $K(1) \cong 1 \amalg 1$ is the classifier for complemented subobjects.

(iii) similarly follows from (i) and 5.4.4(iv): morphisms $I \to K(A \amalg B)$, for an arbitrary I, correspond to K-finite subobjects of $I^*(A \amalg B)$ in $\mathcal{E}/I$, and these correspond in turn to pairs of K-finite subobjects of $I^*(A)$ and $I^*(B)$. □

We note that part (ii) of 5.4.5 justifies the assertion, made earlier in this section, that not all subobjects of K-finite objects need be K-finite. By 5.4.4(iv), we know that complemented subobjects of K-finite objects are always K-finite; however, a general K-finite object may have K-finite subobjects which are not complemented. (We shall return to this point in 5.4.17 below.) Note also that, in conjunction with the fact (5.4.4(ii)) that the support of any K-finite object is K-finite, 5.4.5(ii) tells us that it is true in the internal logic of $\mathcal{E}$ that 'any finite object is either empty or inhabited'. This is characteristic of the fact that many classical modes of reasoning, whilst invalid for arbitrary objects of a topos,

may nevertheless be safely employed for finite objects. As a further example, we sketch the proof of the 'dual Frobenius rule' (cf. 1.3.9).

Lemma 5.4.6 *For any (first-order) formulae ϕ and ψ such that ϕ does not involve the variable y, the sequent*

$$((\forall y)(\phi \vee \psi) \vdash_{\vec{x}} (\phi \vee (\forall y)\psi))$$

is valid in the internal logic of a topos, provided (the object interpreting) the sort of the variable y is a K-finite object.

Proof Let A be the product of (the objects interpreting) the sorts of the variables in $\vec{x}$, and let B be the sort of y. Then ϕ and ψ are respectively interpreted as subobjects $X \rightarrowtail A$ and $Y \rightarrowtail A \times B$. We need to show that $\forall_{\pi_1}((X \times B) \cup Y) \leq X \cup \forall_{\pi_1}(Y)$ in $\mathrm{Sub}(A)$ (in fact we have equality here, since the reverse inclusion is true in general). It is clear that we may form a subobject L of PB which is 'the object of subobjects B' such that $\forall_{\pi_1}((X \times B') \cup (Y \cap (A \times B'))) \leq X \cup \forall_{\pi_1}(Y \cap (A \times B'))$'. Now L clearly contains $\ulcorner 0 \urcorner$ and singletons, and it is closed under binary unions since $\mathrm{Sub}(A)$ is a distributive lattice. Hence $K(B) \leq L$; in particular the top element of PB belongs to L, and so the result is established. □

A similar argument shows that, in a topos where De Morgan's law is satisfied (cf. 4.6.2), the sequent

$$(\neg(\forall y)\phi \vdash_{\vec{x}} (\exists y)\neg\phi)$$

is valid for any ϕ with free variables in the string $\vec{x}, y$, provided the sort of y is K-finite. Indeed, this may be taken to be yet another equivalent way of expressing De Morgan's law in a topos, since it reduces to the usual formulation of the latter when the sort of y is 2.

Although a subobject of a K-finite object need not be K-finite, the following rather technical lemma (due to T. Coquand [249]) provides a partial substitute which is of use in several contexts.

Lemma 5.4.7 *Let A be an object of a topos $\mathcal{E}$, and let A_1, A_2, A_3 be subobjects of A such that $A_1 \leq A_2 \cup A_3$ and A_1 is K-finite. Then it is locally true that A_1 is contained in the union of two K-finite subobjects of A_2 and A_3, i.e. there exists $B \twoheadrightarrow 1$ in $\mathcal{E}$ and K-finite subobjects $A_2' \rightarrowtail B^*A_2$, $A_3' \rightarrowtail B^*A_3$ in $\mathcal{E}/B$ such that $B^*A_1 \leq A_2' \cup A_3'$.*

Classically (i.e. in a Boolean topos), this result would be a triviality, because we could simply take $B = 1$ and $A_j' = A_1 \cap A_j$ $(j = 2, 3)$. However, using 5.4.14 below it is easy to give examples in non-Boolean toposes where we really need to localize, i.e. there do not exist suitable K-finite subobjects with $B = 1$.

Proof It suffices to show that the sentence

$$(\forall u, v, w)(((u \in \ulcorner K(A) \urcorner) \wedge (u \leq v \cup w)) \Rightarrow$$
$$(\exists v')((v' \leq v) \wedge (v' \in \ulcorner K(A) \urcorner) \wedge (u \leq v' \cup w)))$$

(where all variables are of type PA) is valid in the internal language of $\mathcal{E}$, since applying this twice will yield the result in the statement. To do this, we consider the subobject L of PA defined by the term

$$\{u \mid (\forall v, w)((u \leq v \cup w) \Rightarrow (\exists v')((v' \leq v) \wedge (v' \in \ulcorner K(A) \urcorner) \wedge (u \leq v' \cup w)))\},$$

and show that it has the closure properties of 5.4.2. It clearly contains $\ulcorner 0 \urcorner$, since if $u = \ulcorner 0 \urcorner$ then we can take v' to be $\ulcorner 0 \urcorner$ as well. If $u \in L$ and $u' = \{x\} \cup u$, then from $(u' \leq v \cup w)$ we deduce the (local) existence of a finite $v' \leq v$ with $u \leq v' \cup w$, and also $((x \in v) \vee (x \in w))$. If $(x \in v)$ holds, we define $v'' = \{x\} \cup v'$, and if $(x \in w)$ we define $v'' = v'$; in either case, it satisfies the conditions. □

As well as finite coproducts, the class of K-finite objects is also closed under finite products. However, the proof of this fact is a little more complicated: in effect, it involves a 'double induction'.

Proposition 5.4.8 *A product of two K-finite objects is K-finite.*

Proof Let $\theta_{A,B} \colon PA \times B \to P(A \times B)$ be the name of the relation

$$\in_A \times B \rightarrowtail PA \times A \times B \times B \cong PA \times B \times A \times B$$

(where the monomorphism $B \rightarrowtail B \times B$ is the diagonal). In 'elementwise' terms, $\theta_{A,B}$ sends (A', b) to $A' \times \{b\}$; in fact it is a strength for the covariant power-object functor, in the sense of B2.1.4. Since the product of a K-finite object and a singleton is clearly K-finite, $\theta_{A,B}$ restricts to a morphism $K(A) \times B \to K(A \times B)$, which is a strength for the functor K. Hence, if A is K-finite, we have a diagram

$$\begin{tikzcd}
B \arrow[r, "(\top_A, 1_B)"] \arrow[d, tail, "\{\}"] & K(A) \times B \arrow[r, "\theta_{A,B}"] & K(A \times B) \arrow[d, tail] \\
PB \arrow[rr, "P\pi_2"] & & P(A \times B)
\end{tikzcd}$$

which is readily seen to commute, since both ways round name the same subobject of $B \times A \times B$. But $P\pi_2$ is a join-semilattice homomorphism, so the pullback of $K(A \times B)$ along it is a sub-join-semilattice of PB, and we have just seen that

it contains $\{\}\colon B \rightarrowtail PB$; so it contains $K(B)$. Thus if B is also K-finite, we deduce that the composite

$$1 \xrightarrow{\top_B} PB \xrightarrow{P\pi_2} P(A \times B)$$

factors through $K(A \times B)$; but this is precisely the top element of $P(A \times B)$. □

We have seen that K is a functor $\mathcal{E} \to \mathcal{E}$, and the factorization of the singleton map $A \rightarrowtail PA$ through $K(A)$ clearly yields a natural transformation η from the identity functor to K. It is no great surprise that η is the unit of a monad structure on K; what is perhaps more unexpected is that the algebras for this monad turn out to be something very familiar.

Proposition 5.4.9 *For any A, $K(A)$ is the free semilattice generated by A. In particular, the forgetful functor from the category* **SLat**$(\mathcal{E})$ *of internal semilattices in $\mathcal{E}$ to $\mathcal{E}$ is monadic, even if $\mathcal{E}$ does not have a natural number object.*

Proof By definition, $K(A)$ carries a semilattice structure, and $K(f)$ is a semilattice homomorphism for each morphism $f\colon A \to B$ of $\mathcal{E}$, since it is a restriction of the semilattice homomorphism $\exists f\colon PA \to PB$. So we can regard K as a functor $\mathcal{E} \to$ **SLat**$(\mathcal{E})$.

Now let B be an internal semilattice in $\mathcal{E}$. The equalizer of the semilattice operation $\vee\colon B \times B \to B$ and the second projection is a partial order on B; we may construct the object $C \rightarrowtail PB$ of subobjects of B having least upper bounds for this ordering, and since least upper bounds are unique when they exist, they define a morphism $C \to B$. We claim that $K(B) \leq C$, so that this restricts to a morphism $\epsilon_B\colon K(B) \to B$. Clearly, C is a sub-join-semilattice of PB, since the zero element of B is a least upper bound for the zero subobject, and the least upper bound of the union of two members of C is the join of their least upper bounds; but the least upper bound of a singleton is its unique member, and so $\{\}\colon B \to PB$ factors through C. Thus the claim is established; moreover, it follows from the proof that ϵ_B is a semilattice homomorphism.

Now suppose we are given an arbitrary morphism $f\colon A \to B$, where B has a semilattice structure. Then the composite

$$K(A) \xrightarrow{K(f)} K(B) \xrightarrow{\epsilon_B} B$$

is a semilattice homomorphism; and its composite with η_A is f since $\epsilon_B\eta_B$ is the identity on B. But if we had two semilattice homomorphisms $K(A) \rightrightarrows B$ extending f, their equalizer would be a sub-semilattice of $K(A)$ containing the image of η_A, and hence would be the whole of $K(A)$. Thus the adjunction is established.

Finally, the monadicity of the forgetful functor **SLat**$(\mathcal{E}) \to \mathcal{E}$ follows from 5.3.1. □

We may use 5.4.9 to deduce further properties of K-finiteness from known results about finitary algebraic theories. For example, it follows immediately from 5.4.9 and 5.3.7 that inverse image functors commute up to isomorphism with the functor K. But in fact a stronger result is true: arbitrary bicartesian (that is, cartesian and cocartesian) functors between toposes commute with K, even if they do not have right adjoints. Before proving this, we need an alternative universal characterization of $K(A)$, which stands in the same relation to 5.4.9 as list objects do to free monoids.

Given an object A of a topos $\mathcal{E}$, let us write $\mathbb{T}_A$ for the internal algebraic theory freely generated by one nullary operation and an A-indexed family of unary operations, so that a $\mathbb{T}_A$-algebra is simply an object B equipped with morphisms $b\colon 1 \to B$ and $t\colon A \times B \to B$. (Thus a list object LA, if it exists, is an initial object of the category $\mathbb{T}_A$-**Mod**$(\mathcal{E})$.) We shall say that a $\mathbb{T}_A$-algebra (B, b, t) is *idempotent* (resp. *commutative*) if it satisfies the sequent $(\top \vdash_{x,y} (t(x, t(x, y)) = t(x, y)))$ (resp.

$$(\top \vdash_{x,x',y} (t(x, t(x', y)) = t(x', t(x, y))));$$

and we shall write $\mathbb{S}_A$ for the theory of idempotent commutative $\mathbb{T}_A$-algebras. Clearly, the structure maps $\ulcorner 0 \urcorner\colon 1 \to K(A)$ and $\sigma_A\colon A \times K(A) \to K(A)$ of 5.4.2 make $K(A)$ into an $\mathbb{S}_A$-algebra.

Proposition 5.4.10 *For any A, $(K(A), \ulcorner 0 \urcorner, \sigma_A)$ is an initial object of the category $\mathbb{S}_A$-**Mod**$(\mathcal{E})$.*

Proof Suppose given an arbitrary $\mathbb{S}_A$-algebra (B, b, t). Consider the transpose $\bar{t}\colon A \to B^B$ of t; let $\overline{A} \rightarrowtail B^B$ denote its image. Since (B, b, t) is idempotent, it is clear that $\overline{A}$ is contained in the subobject $\mathrm{Idem}(B) \rightarrowtail B^B$ of idempotent endomorphisms of B, defined by the term

$$\{f\colon B^B \mid (f \circ f = f)\}$$

(where we are using $f \circ g$ as a shorthand for the term $\lambda x \,.\, \mathsf{app}(f, \mathsf{app}(g, x)))$. We define successively smaller subobjects T, S of B^B by the terms

$$\{f\colon B^B \mid ((f \in \ulcorner \mathrm{Idem}(B) \urcorner) \wedge (\forall g)((g \in \ulcorner \overline{A} \urcorner) \Rightarrow (f \circ g = g \circ f)))\}$$

and

$$\{f\colon B^B \mid ((f \in \ulcorner T \urcorner) \wedge (\forall g)((g \in \ulcorner T \urcorner) \Rightarrow (f \circ g = g \circ f)))\}$$

respectively. Since (B, b, t) is commutative, it is easy to see that we have $\overline{A} \leq S$ in $\mathrm{Sub}(B^B)$; but since the members of S commute with each other they are closed under composition, i.e. S is a sub-monoid of B^B. By construction, it is also idempotent and commutative, so it is a semilattice. Hence by 5.4.9 we may

extend the composite $\bar{t}\colon A \twoheadrightarrow \overline{A} \rightarrowtail S$ uniquely to a semilattice homomorphism $\hat{t}\colon K(A) \to S$. Now the composite

$$K(A) \xrightarrow{\hat{t}} S \rightarrowtail B^B \xrightarrow{1 \times b} B^B \times B \xrightarrow{\mathrm{ev}} B$$

is easily seen to be a homomorphism of $\mathbb{S}_A$-algebras; and it is unique, since the equalizer of two such homomorphisms would be a sub-$\mathbb{S}_A$-algebra of $K(A)$, and hence would be the whole of $K(A)$ by 5.4.2. □

In fact $\mathbb{S}_A$-**Mod**$(\mathcal{E})$, like **SLat**$(\mathcal{E})$, is monadic over $\mathcal{E}$ even if $\mathcal{E}$ does not have a natural number object; the free $\mathbb{S}_A$-algebra on an arbitrary object B is simply $K(A) \times (1 \amalg B)$. However, we shall not need this; instead, we turn to a characterization of $K(A)$ which is similar to Freyd's characterization of natural number objects and list objects:

Corollary 5.4.11 *Let A and B be objects of a topos. Then $B \cong K(A)$ iff there exist morphisms $b\colon 1 \to B$ and $t\colon A \times B \to B$ for which the coherent sequents*

$$\begin{gathered}(\top \vdash_{x,y} (t(x,t(x,y)) = t(x,y)))\\ (\top \vdash_{x,x',y} (t(x,t(x',y)) = t(x',t(x,y))))\\ ((t(x,y) = b) \vdash_{x,y} \bot)\\ ((t(x,t(x',y)) = t(x',y)) \vdash_{x,x',y} ((x = x') \vee (t(x,y) = y)))\end{gathered}$$

and

$$(\top \vdash_y ((y = b) \vee (\exists x)(t(x,y) = y)))$$

are satisfied (where x and y are variables of types A and B respectively), and the diagram

$$A \times B \underset{\pi_2}{\overset{t}{\rightrightarrows}} B \longrightarrow 1$$

is a coequalizer.

Proof If $B = K(A)$, then $\ulcorner 0 \urcorner\colon 1 \to B$ and the $\sigma_A\colon A \times B \to B$ are easily seen to satisfy the given conditions. In particular, the last displayed sequent expresses the fact, which we noted after 5.4.5, that any K-finite object is either empty or inhabited. To verify the last condition, let $R \rightrightarrows B$ be the equivalence relation generated by the parallel pair (σ_A, π_2) (i.e. the kernel-pair of its coequalizer); then it is easy to see that the R-equivalence class of $\ulcorner 0 \urcorner$ satisfies the closure conditions of 5.4.2, and is thus the whole of $K(A)$.

Conversely, suppose B satisfies the conditions. The first two say that (B, b, t) is an $\mathbb{S}_A$-algebra, so by 5.4.10 we have a unique homomorphism $f\colon K(A) \to B$. But the last two conditions imply that B satisfies the analogue of Peano's fifth postulate: that is, any subobject $B' \rightarrowtail B$ which is a sub-$\mathbb{S}_A$-algebra must be the whole of B. (The proof of this is essentially the same as the proof

of (ii) $\Rightarrow$ (iii) in 5.1.2.) Now we define a morphism $g\colon B \to PA$ by the term-in-context $y\,.\{x\colon A \mid (t(x,y) = y)\}$. Clearly, we have $(\top \vdash (g(b) = \ulcorner 0 \urcorner))$ and $(\top \vdash_{x,y} (g(t(x,y)) = \{x\} \cup g(y)))$ by the third and fourth displayed sequents, so the subobject $B' \rightarrowtail B$ defined by the term $\{y\colon B \mid (g(y) \in \ulcorner K(A) \urcorner)\}$ is a sub-$\mathbb{S}_A$-algebra, and hence we may regard g as a morphism $B \to K(A)$. Moreover, it is clearly an $\mathbb{S}_A$-algebra homomorphism; and the composite fg (resp. gf) is equal to the identity, since in each case the equalizer of the composite and the identity is a sub-$\mathbb{S}_A$-algebra of B (resp. $K(A)$). □

Corollary 5.4.12 *Let $F\colon \mathcal{E} \to \mathcal{F}$ be a bicartesian functor (for example, an inverse image functor). If A is a K-finite object of $\mathcal{E}$, then $F(A)$ is K-finite in $\mathcal{F}$.*

Proof By 5.4.11 (or by 5.4.9 and 5.3.7, if F is an inverse image functor), we have an isomorphism $F(K(A)) \cong K(F(A))$; and this is easily seen to be an isomorphism of internal posets. But $F(K(A))$ clearly has a greatest element if $K(A)$ has. □

An alternative proof of 5.4.12, in the case when the toposes $\mathcal{E}$ and $\mathcal{F}$ have natural number objects, may be deduced from A2.5.7 and the fact (3.2.10) that K-finite objects are exactly the models of a σ-coherent theory $\mathbb{K}$; but the latter identification depends on 5.4.13 below, as we remarked in 1.2.15(k).

The converse of 5.4.12 does not hold even if F is conservative. It follows from 5.4.12 that if A is a K-finite object in a functor category $[\mathcal{C}, \mathbf{Set}]$, then $A(c)$ is a finite set for each object c of $\mathcal{C}$; but this condition is not sufficient for K-finiteness of A, as we shall see in 5.4.14 below.

An important consequence of 5.4.9 is that it enables us to determine the precise relationship between K-finiteness and cardinal-finiteness in the case when $\mathcal{E}$ has a natural number object.

Theorem 5.4.13 *Let $\mathcal{E}$ be a topos with a natural number object, A an object of $\mathcal{E}$. Then A is K-finite iff it is locally a quotient of a finite cardinal, i.e. iff there exist $B \twoheadrightarrow 1$ in $\mathcal{E}$, a natural number p in $\mathcal{E}/B$ and an epimorphism $[p] \twoheadrightarrow B^*(A)$.*

Proof First we observe that finite cardinals are K-finite: this is an immediate consequence of 5.2.1 and 5.4.4(i) and (iii). Hence, by 5.4.4(ii), any quotient of a finite cardinal is K-finite; and since K-finiteness is equivalent to the validity of a sentence in the internal language of $\mathcal{E}$, it is reflected by conservative logical functors such as B^*. This proves the right-to-left implication.

For the converse, we observe that, since the algebraic theory of semilattices is a quotient of that of monoids (or, if you prefer, because $\mathbb{S}_A$ is a quotient of $\mathbb{T}_A$), the free semilattice $K(A)$ on an object A must be a quotient of the free monoid $L(A) = \Sigma_N(N^*(A)^C)$ constructed in 5.3.3. Since the epimorphism $L(A) \twoheadrightarrow K(A)$ is a monoid homomorphism, it is easy to see that its 'elementwise' description must be as the operation which sends a morphism $[p] \to A$ to its image considered as a K-finite subobject of A. In particular, if A is K-finite and

we form the pullback

$$
\begin{tikzcd}
B \arrow[r, two heads] \arrow[d] & 1 \arrow[d, "\top_A"] \\
L(A) \arrow[r, two heads] & K(A)
\end{tikzcd}
,
$$

then the natural number p in $\mathcal{E}/B$ and the morphism $[p] \to B^*(A)$ which correspond to the left vertical map in this diagram have the required properties. □

Example 5.4.14 Let $\mathcal{C}$ be a small category. Then an object A of the functor category $[\mathcal{C}, \mathbf{Set}]$ is K-finite iff each $A(c)$, $c \in \mathrm{ob}\,\mathcal{C}$, is a finite set, and each $A(f)$, $f: c \to d$ in $\mathcal{C}$, is surjective. For we saw in 5.2.8(a) that the finite cardinals in $[\mathcal{C}, \mathbf{Set}]$ are finite-set-valued functors which are constant on each connected component of $\mathcal{C}$. Clearly, if A is a quotient of such a functor, then all the $A(f)$ must be surjective; and this property is 'reflected by localizations', so the necessity of the condition is established. For the sufficiency, observe that if the condition holds then, for each object c of $\mathcal{C}$, $H_c^*(A)$ is a quotient of $H_c^*([n])$ in $[\mathcal{C}, \mathbf{Set}]/H_c$, where n is the number of elements in $A(c)$ (and H_c denotes the functor $\mathcal{C}(c, -)$); and the coproduct of all the H_c has global support.

Remark 5.4.15 In contrast to finite cardinals, K-finite objects are not internally projective in general; nor is exponentiation to the power of a K-finite object preserved by inverse image functors (cf. 5.2.9(ii) and 5.2.11). For a counterexample to both, let $\mathcal{E} = [\mathbf{2}, \mathbf{Set}]$ be the Sierpiński topos, and consider the object $A = (2 \to 1)$, where n denotes an n-element set. It is clear that we have an epimorphism $[sso] \twoheadrightarrow A$, and hence that A is K-finite; but $(-)^A$ does not preserve this epimorphism, since $[sso]^A$ and A^A respectively look like $(2 \to 2)$ and $(4 \to 1)$. Also, if $c: \mathbf{Set} \to [\mathbf{2}, \mathbf{Set}]$ is the closed point of $\mathcal{E}$, then c^* fails to preserve the exponential $[sso]^A$. It is true in the Sierpiński topos that B^A is K-finite provided both A and B are (and it is true in any topos with a natural number object that an object of the form $B^{[p]}$ is K-finite if B is, by an easy induction using 5.4.8), but in general the class of K-finite objects can fail to be closed under exponentials: for a counterexample, take $\mathcal{E}$ to be the topos of diagrams of shape $(\bullet \leftarrow \bullet \rightarrow \bullet)$ in $\mathbf{Set}$, and consider the exponential $[sso]^A$, where A is the K-finite object $(1 \leftarrow 2 \rightarrow 2)$.

In general, the K-finiteness of an object A neither implies nor is implied by that of its power object PA. It is easy to give examples of toposes in which $\Omega \cong P(1)$ is not K-finite: one such is the functor category whose subobject classifier is displayed before A1.6.10. On the other hand, the nontrivial subterminal object $U = (0 \to 1)$ in the Sierpiński topos $[\mathbf{2}, \mathbf{Set}]$ is not complemented and hence not K-finite, but PU is isomorphic to $1 \amalg 1$ and hence K-finite. In contrast, we have:

Proposition 5.4.16 *An object A is K-finite iff $K(A)$ is K-finite.*

Proof First suppose A is K-finite. We may form the subobject $L \rightarrowtail PA$ of subobjects $A' \rightarrowtail A$ such that $K(A')$ is finite, as the interpretation of an appropriate formula in the internal language of $\mathcal{E}$. Our aim is to show that L is a sub-join-semilattice of PA and contains $\{\}\colon A \rightarrowtail PA$. The proof that it contains 0 and singletons is easy, since we know that $K(0) \cong 1$ and $K(1) \cong 1 \amalg 1$ are both K-finite. So suppose we have subobjects A' and A'' such that $K(A')$ and $K(A'')$ are both K-finite. Then the epimorphism $A' \amalg A'' \twoheadrightarrow A' \cup A''$ induces an epimorphism $K(A' \amalg A'') \twoheadrightarrow K(A' \cup A'')$, by 5.4.3(ii); but $K(A' \amalg A'') \cong K(A') \times K(A'')$ by 5.4.5(iii), and so it is K-finite by 5.4.8. Hence $K(A' \cup A'')$ is K-finite by 5.4.4(ii). Thus L contains $K(A)$, and in particular $\top_A$ factors through it; i.e. $K(A)$ is K-finite.

Conversely, suppose $K(A)$ is K-finite. We note that the counit map $\epsilon_{K(A)}\colon K(K(A)) \to K(A)$ is order-preserving (since it is a semilattice homomorphism) and (split) epimorphic; so it must map the greatest element of $K(K(A))$ to a greatest element for $K(A)$. So A is K-finite. □

The operation $A \mapsto K(A)$ thus acts something like a 'power-object functor' for the subcategory $\mathcal{E}_{Kf}$ of K-finite objects of $\mathcal{E}$. Of course, it does not in general have all the properties of the power-object functor in a topos: in particular, it does not normally carry the structure of a contravariant functor, since the pullback of a K-finite subobject $B' \rightarrowtail B$ along a map $A \to B$ need not be K-finite, even if A and B both are. However, if we restrict our attention to objects which are decidable (i.e. have complemented diagonal) as well as K-finite, this problem can be overcome:

Lemma 5.4.17 *If A is decidable, then every K-finite subobject of A is complemented; in fact $K(A) \leq 2^A$ as subobjects of PA.*

Proof 2^A is a sub-join-semilattice (in fact a sublattice) of PA, since finite unions of complemented subobjects are complemented. The assertion that A is decidable says precisely that $\{\}\colon A \to PA$ factors through 2^A, so the second assertion of the lemma is immediate from the definition of $K(A)$. The first follows from applying the second to global elements of PA. □

If A is K-finite, then applying the proof of 5.4.4(iv) to the generic complemented subobject of A in $\mathcal{E}/(2^A)$ yields the information that $2^A \leq K(A)$ in $\mathrm{Sub}(PA)$; so if A is both decidable and K-finite then $K(A)$ coincides with 2^A. We may thus conclude:

Theorem 5.4.18 *For any topos $\mathcal{E}$, the full subcategory $\mathcal{E}_{dKf}$ of decidable K-finite objects of $\mathcal{E}$ is a Boolean topos. In particular, if $\mathcal{E}$ itself is Boolean, then $\mathcal{E}_{Kf}$ is a topos.*

Proof A product of two K-finite objects is K-finite, as we saw in 5.4.8; and a product of two decidable objects is decidable, by A1.4.15. Since 1 is both decidable and K-finite, it follows that $\mathcal{E}_{dKf}$ is closed under finite products in $\mathcal{E}$.

It is also closed under equalizers, since the equalizer of two morphisms $A \rightrightarrows B$ is a complemented subobject of A if B is decidable (and is therefore K-finite if A is), and any subobject of a decidable object is decidable. Thus $\mathcal{E}_{dKf}$ is cartesian. Also, monomorphisms in $\mathcal{E}_{dKf}$ coincide with those in $\mathcal{E}$; so every subobject in $\mathcal{E}_{dKf}$ is complemented, and hence $K(A) \cong 2^A$ has the universal property of a power object for A in $\mathcal{E}_{dKf}$, provided it lies in this subcategory. We already know it is K-finite, by 5.4.16; so we need only prove it is decidable.

For this, we argue inductively as usual to prove that, for any K-finite A, 2^A is decidable. We may form the subobject $L \rightarrowtail PA$ of subobjects $A' \rightarrowtail A$ for which $2^{A'}$ is decidable; this clearly contains 0 and singletons since $2^0 \cong 1$ and 2 are decidable, and it is closed under binary unions since $2^{A' \cup A''}$ admits a monomorphism to $2^{A' \amalg A''} \cong 2^{A'} \times 2^{A''}$, corresponding to the epimorphism $A' \amalg A'' \twoheadrightarrow A' \cup A''$. So $K(A) \leq L$, and hence if A itself is K-finite $\top_A$ factors through L.

This completes the proof that $\mathcal{E}_{dKf}$ is (a weak topos, and hence by A2.3.4) a topos. The second assertion follows since in a Boolean topos every object is decidable. □

Example 5.4.19 Let $\mathcal{C}$ be a small category. Combining 5.4.14 and A1.4.16, we see that a functor $A: \mathcal{C} \to \mathbf{Set}$ is decidable and K-finite as an object of $[\mathcal{C}, \mathbf{Set}]$ iff $A(c)$ is finite for each object c of $\mathcal{C}$ and $A(f)$ is bijective for each morphism f of $\mathcal{C}$. Thus we see that the topos $[\mathcal{C}, \mathbf{Set}]_{dKf}$ is equivalent to the topos $[\mathcal{G}, \mathbf{Set}_f]$ (cf. A2.1.5), where $\mathcal{G}$ is the groupoid reflection of $\mathcal{C}$, i.e. the category obtained by freely adjoining inverses for all the morphisms of $\mathcal{C}$. In particular, if $\mathcal{C}$ itself is a groupoid, then $[\mathcal{C}, \mathbf{Set}]_{dKf} \cong [\mathcal{C}, \mathbf{Set}_f]$.

More generally, by combining 5.4.13 and 5.2.6 we see that, in a topos with a natural number object, an object is decidable and K-finite iff it is locally isomorphic to a finite cardinal.

Remark 5.4.20 It is possible for $\mathcal{E}_{Kf}$ to be a topos even if $\mathcal{E}$ is not Boolean, but (perhaps surprisingly) this cannot happen unless it coincides with $\mathcal{E}_{dKf}$. In fact it can be shown that *the following conditions on a topos $\mathcal{E}$ are equivalent:*

(i) *$\mathcal{E}_{Kf}$ is a topos.*

(ii) *Every K-finite object of $\mathcal{E}$ is decidable.*

(iii) *Every subobject of a K-finite object of $\mathcal{E}$ is K-finite.*

(iv) *The (external) lattice* $\mathrm{Sub}_{\mathcal{E}}(1)$ *of subterminal objects of $\mathcal{E}$ is Boolean.*

(v) *$\mathcal{E}_{Kf}$ is closed under the formation of equalizers in $\mathcal{E}$.*

We omit the proof, which can be found in [543].

The fact that subobjects of K-finite objects are not in general K-finite may be 'remedied' by modifying the definition in either of two ways. We could simply consider the class of objects which occur as subobjects of K-finite objects (cf. 5.4.20(iii)); however, this yields a condition which is not local (i.e. an object

A of $\mathcal{E}$ may be such that B^*A is a subobject of a K-finite object of $\mathcal{E}/B$ for some $B \twoheadrightarrow 1$, without A being a subobject of any K-finite object of $\mathcal{E}$). A better approach is to replace the subobject $\{\}\colon A \rightarrowtail PA$ in the definition of $K(A)$ by the partial-map classifier $\tilde{A}$, regarded (as in A2.4.10) as 'the object of subobjects of A having at most one element'.

Formally, we define $\tilde{K}(A)$ to be the sub-join-semilattice of PA generated by $\tilde{A} \rightarrowtail PA$.

Lemma 5.4.21 *$\tilde{K}(A)$ is downwards-closed as a subobject of PA.*

Proof For any subobject $B \rightarrowtail A$, the mapping $(-) \cap B\colon PA \to PA$ is a join-semilattice homomorphism (since PA is a distributive lattice) and it maps $\tilde{A}$ into itself since $\tilde{A}$ is downwards-closed. So it maps $\tilde{K}(A)$ into itself. Since this is true for all B, $\tilde{K}(A)$ is downwards-closed. □

Thus we see that, for a given A, the top element $\ulcorner A \urcorner$ of PA factors through $\tilde{K}(A)$ iff the latter is the whole of PA; we say that A is *$\tilde{K}$-finite* if these conditions hold. The elementary properties of the functor $\tilde{K}$, and of $\tilde{K}$-finite objects, may be developed along the same lines as those of K in the earlier part of this section; but it is quicker to use

Proposition 5.4.22 *An object A is $\tilde{K}$-finite iff it is locally a subobject of a K-finite object.*

Proof Since $\tilde{K}(A)$ is generated by $\tilde{A}$ as a semilattice, it must be an epimorphic image of the free semilattice $K(\tilde{A})$. It is easy to see that in 'elementwise' terms the epimorphism $K(\tilde{A}) \twoheadrightarrow \tilde{K}(A)$ sends a K-finite subobject $B \rightarrowtail \tilde{A}$ to the subobject $B' \rightarrowtail A$ defined by the pullback

$$\begin{tikzcd} B' \arrow[r, tail] \arrow[d, tail] & A \arrow[d, tail, "i"] \\ B \arrow[r, tail] & \tilde{A} \end{tikzcd} \quad ;$$

for this operation is a join-semilattice homomorphism, and has the right effect on the generators. So, if $\ulcorner A \urcorner\colon 1 \to PA$ factors through $\tilde{K}(A) \rightarrowtail PA$, then by forming the pullback

$$\begin{tikzcd} C \arrow[r] \arrow[d, two heads] & K(\tilde{A}) \arrow[d, two heads] \\ 1 \arrow[r, "\ulcorner A \urcorner"] & \tilde{K}(A) \end{tikzcd}$$

we obtain a K-finite subobject B of $C^*(\tilde{A})$ for which the corresponding B' is the whole of $C^*(A)$. Conversely, if A is a subobject of a K-finite object B, then the

image of the map $B \to \tilde{A}$ corresponding to the partial map $(A \rightarrowtail B, 1_A)$ defines an element of $K(\tilde{A})$ whose image under $K(\tilde{A}) \twoheadrightarrow \tilde{K}(A)$ is $\ulcorner A \urcorner$. And since $\tilde{K}$-finiteness is a local property by the way in which it is defined, we may conclude that any object which is locally a subobject of a K-finite object is $\tilde{K}$-finite. $\square$

In particular, if $\mathcal{E}$ has a natural number object, then an object of $\mathcal{E}$ is $\tilde{K}$-finite iff it is locally a subquotient of a finite cardinal, by 5.4.13 and 5.4.22.

$\tilde{K}$-finiteness, like K-finiteness and cardinal-finiteness, corresponds to an induction principle: if P is a property of subobjects of A (expressible by a formula in the internal language of $\mathcal{E}$) such that P holds for all subsingletons (that is, all subobjects $U \rightarrowtail A$ whose domains are subterminal objects), and P holds for a union $B_1 \cup B_2$ provided it holds for B_1 and B_2, then P holds for all $\tilde{K}$-finite subobjects of A. In practice, this induction principle is less useful than that associated with K-finiteness, since it is hard to 'get the induction started'; i.e. the class of properties which hold for all subsingletons is uncomfortably restricted. However, we note one case where it may be applied:

Proposition 5.4.23 *The category $\mathcal{E}_{\tilde{K}f}$ of $\tilde{K}$-finite objects of $\mathcal{E}$ is a logical subtopos of $\mathcal{E}$ iff Ω is K-finite.*

Proof Clearly, if $\mathcal{E}_{\tilde{K}f}$ is a logical subtopos then it must contain Ω; but Ω is injective by A2.2.6, so if it is $\tilde{K}$-finite then it is (locally, and hence globally) K-finite.

Conversely, suppose Ω is K-finite. It follows easily from 5.4.22 that $\mathcal{E}_{\tilde{K}f}$ is closed under finite limits (and finite colimits) in $\mathcal{E}$; so we need only prove that a power object PA is $\tilde{K}$-finite (equivalently, K-finite) if A is $\tilde{K}$-finite. For this we use the induction principle mentioned above: consider the sentence which asserts about a subobject B of A that PB is K-finite. It holds for subsingletons, since if B is subterminal then PB is a retract of $P1 = \Omega$, by A2.2.5. And if $B = B_1 \cup B_2$, then PB is a retract of $P(B_1 \amalg B_2) \cong PB_1 \times PB_2$, so the result follows from 5.4.8. $\square$

Corollary 5.4.24 *If $\mathcal{E}$ is a topos whose subterminal objects form a generating family (for example, a localic* **Set**-*topos), then $\mathcal{E}_{\tilde{K}f}$ is a topos iff Ω is K-finite in $\mathcal{E}$.*

Proof One direction follows from 5.4.23. Conversely, suppose $\mathcal{E}_{\tilde{K}f}$ is a topos; let Ω' denote its subobject classifier. The generic monomorphism in $\mathcal{E}_{\tilde{K}f}$ is a monomorphism in $\mathcal{E}$, so it is classified by a morphism $\Omega' \to \Omega$; but since every subterminal object of $\mathcal{E}$ is in $\mathcal{E}_{\tilde{K}f}$, it follows easily from the hypothesis that this morphism is an isomorphism. So Ω is $\tilde{K}$-finite in $\mathcal{E}$; as in the proof of 5.4.23, this implies that it is K-finite. $\square$

In general, however, $\mathcal{E}_{\tilde{K}f}$ may be a topos without being a logical subtopos of $\mathcal{E}$. Examples may be constructed using 5.4.20 above, or alternatively using

the following:

Example 5.4.25 If $\mathcal{E}$ is a functor category $[\mathcal{C}, \mathbf{Set}]$, then a functor $A\colon \mathcal{C} \to \mathbf{Set}$ which is $\tilde{K}$-finite as an object of $\mathcal{E}$ must take finite sets as values, though its transition maps need not be either monic or epic. However, in general not every finite-set-valued functor on $\mathcal{C}$ is $\tilde{K}$-finite as an object of $[\mathcal{C}, \mathbf{Set}]$, even if $\mathcal{C}$ itself is finite. For if $\mathcal{C}$ contains a parallel pair (f, g) such that $f \neq g$ but $fh = gh$ for some h, then it is clear that any K-finite object A of $[\mathcal{C}, \mathbf{Set}]$ must satisfy $A(f) = A(g)$; so this condition is inherited by any subobject of a K-finite object, and also by any $\tilde{K}$-finite object (the latter because, if B is a well-supported object such that $B^*(A)$ is a subobject of a K-finite object, then $B(c)$ must be nonempty, where c is the domain of h). For finite categories $\mathcal{C}$, this is the only obstruction to $\tilde{K}$-finiteness: *if $\mathcal{C}$ is a finite category, then $[\mathcal{C}, \mathbf{Set}]_{\tilde{K}f}$ may be identified with $[\mathcal{C}/R, \mathbf{Set}_f]$, where R is the congruence on $\mathcal{C}$ generated by all pairs (f, g) for which there exists h with $fh = gh$.* In particular, $[\mathcal{C}, \mathbf{Set}]_{\tilde{K}f}$ is a topos for any finite $\mathcal{C}$, although if the congruence R is nontrivial it need not be a logical subtopos of $[\mathcal{C}, \mathbf{Set}]$.

To prove this, consider first the case when $\mathcal{C}$ has an initial object (so that R consists of all parallel pairs in $\mathcal{C}$). If $A\colon \mathcal{C} \to \mathbf{Set}_f$ is a functor respecting the congruence R, we define

$$\overline{A}(c) = \left(\coprod_{d \in \mathrm{ob}\ \mathcal{C}} A(d) \right) \Big/ \equiv_c$$

for each $c \in \mathrm{ob}\ \mathcal{C}$, where $\equiv_c$ is the equivalence relation which identifies $x \in A(d)$ with $y \in A(e)$ iff there exist $f\colon d \to c$ and $g\colon e \to c$ such that $A(f)(x) = A(g)(y)$. (The fact that A respects R ensures that this is indeed an equivalence relation.) It is easy to see that if there exists a morphism $c \to d$ in $\mathcal{C}$ then $\equiv_c \subseteq \equiv_d$; so the assignment $c \mapsto \overline{A}(c)$ may be made into a functor $\mathcal{C} \to \mathbf{Set}_f$ whose transition maps are surjective, and which is therefore K-finite by 5.4.14. Moreover, $\equiv_c$ cannot identify distinct elements of $A(c)$, so we have an injective map $A(c) \rightarrowtail \overline{A}(c)$, which is easily seen to be the c-component of a natural transformation $A \to \overline{A}$. So we have expressed A as a subobject of a K-finite object of $[\mathcal{C}, \mathbf{Set}]$.

For the general case, let c be an arbitrary object of $\mathcal{C}$ and let H_c denote the functor $\mathcal{C}(c, -)$. The total category of the discrete opfibration corresponding to H_c (as in A1.1.7) is the co-slice category $c \backslash \mathcal{C}$, which has an initial object; moreover, if a functor A respects the congruence R on $\mathcal{C}$, then $H_c^*(A)$, regarded as a functor on $c\backslash\mathcal{C}$, respects the corresponding congruence on the latter. So if A also takes finite sets as values, we see that $H_c^*(A)$ is a subobject of a K-finite object of $[\mathcal{C}, \mathbf{Set}]/H_c$; and since the coproduct of all the H_c, $c \in \mathrm{ob}\ \mathcal{C}$, is well-supported it follows that A is locally a subobject of a K-finite object. So the result follows from 5.4.22.

Note in particular that if every morphism of $\mathcal{C}$ is epic, then the congruence R is trivial, and so $[\mathcal{C}, \mathbf{Set}]_{\tilde{K}f} \cong [\mathcal{C}, \mathbf{Set}_f]$ is a logical subtopos of $[\mathcal{C}, \mathbf{Set}]$. (Of

course, it is easy to see that in this case Ω has surjective transition maps, and is therefore K-finite.) However, the reader should be warned that the quotient category $\mathcal{C}/R$ may not have the property that all its morphisms are epic; it may be necessary to iterate the quotient construction a finite number of times before we arrive at a category $\mathcal{C}/\overline{R}$ with the latter property. Moreover, the objects of $[\mathcal{C}, \mathbf{Set}]$ which are globally expressible as subobjects of K-finite objects are exactly the finite-set-valued functors which respect this larger congruence $\overline{R}$; so, as we claimed earlier, the latter property is not equivalent to $\tilde{K}$-finiteness. For more details, see [543].

For an infinite category $\mathcal{C}$, there are further restrictions on the objects of $[\mathcal{C}, \mathbf{Set}]_{\tilde{K}f}$. For example, if $\mathcal{C}$ is the ordered set of natural numbers, then it is easy to see that the functor A defined by $A(n) = \{0,1\}$ for each n, with each transition map $A(n) \to A(n+1)$ sending both 0 and 1 to 0, is not $\tilde{K}$-finite. Moreover, $[\mathcal{C}, \mathbf{Set}]_{\tilde{K}f}$ is not a topos in this case, since it lacks a subobject classifier. (Recall that we pictured the subobject classifier of $[\mathcal{C}, \mathbf{Set}]$, for this particular $\mathcal{C}$, before A1.6.10. The object A which we have just defined occurs naturally as a subobject of Ω; in fact it is the Higgs object of the topos $[\mathcal{C}, \mathbf{Set}]$, as defined before 4.5.10.)

In 5.2.14 and 5.2.15 we saw one way in which preservation of finite colimits may be broken down into preservation of finite colimits and of filtered colimits. The notion of K-finiteness gives us another way of doing this, which is simpler in that it does not require the topos to have a natural number object; however, it works only for indexed preorders. The reason is that, if we are dealing with an indexed category $\mathbb{C}$ which is not a preorder, then an object A in a fibre $\mathcal{C}^J$ of $\mathbb{C}$ will not in general satisfy $A \cong A + A$; so, if we are given a morphism $x\colon [p] \to J$ to J from a finite cardinal $[p]$, with (K-finite) image $m\colon J' \rightarrowtail J$, then the coproducts $\Sigma_{[p]} x^* A$ and $\Sigma_{J'} m^* A$ will in general be different. But if $\mathbb{C}$ is an indexed preorder, then they coincide; so we can forget about the $L(J)$-indexed family of all morphisms from finite cardinals to J, and work instead with the poset $K(J)$ of K-finite subobjects of J. Thus we have

Proposition 5.4.26 *Let $\mathcal{S}$ be a topos, $\mathbb{C}$ and $\mathbb{D}$ two cocomplete $\mathcal{S}$-indexed preorders which are stacks for the coherent coverage on $\mathcal{S}$, and $F\colon \mathbb{C} \to \mathbb{D}$ an $\mathcal{S}$-indexed functor.*

(i) *If F preserves (fibrewise) finite joins, then it preserves joins indexed by K-finite objects of $\mathcal{S}$, i.e. the diagram*

$$\begin{array}{ccc} \mathcal{C}^I & \xrightarrow{F^I} & \mathcal{D}^I \\ \Big\downarrow{\scriptstyle \Sigma_x} & & \Big\downarrow{\scriptstyle \Sigma_x} \\ \mathcal{C}^J & \xrightarrow{F^J} & \mathcal{D}^J \end{array}$$

commutes up to isomorphism whenever $x: I \to J$ *is a* K*-finite object of* $\mathcal{S}/J$.

(ii) *If* F *preserves finite joins and* $\mathcal{S}$*-indexed directed joins, then it is* $\mathcal{S}$*-cocontinuous.*

Proof (i) For simplicity, we consider the case $J = 1$. Given an object A of $\mathcal{C}^I$, we may construct in $\mathcal{S}$ the subobject $L \rightarrowtail PI$ indexing those subobjects $I' \rightarrowtail I$ such that the square commutes for A restricted to $\mathcal{C}^{I'}$; the fact that $\mathbb{C}$ and $\mathbb{D}$ are stacks and F preserves fibrewise finite joins easily implies that L is a sub-join-semilattice of PI, and it contains singletons since the vertical edges of the diagram reduce to identity maps if I' is a singleton. So $K(I) \leq L$, and in particular the top element $\ulcorner I \urcorner$ of PI factors through L.

(ii) Again, we deal with the case when $J = 1$. For any I, $K(I)$ is a directed internal poset in $\mathcal{S}$, because it is a join-semilattice; and we may compute the join of an I-indexed family $A \in \text{ob}\, \mathcal{C}^I$ by first computing the joins of all its K-finite subfamilies, and then forming the directed join of the latter. So if F commutes with joins of both types, then it commutes with arbitrary joins; but this suffices for cocontinuity, since we are dealing with preorders. □

Suggestions for further reading: Acuña-Ortega & Linton [14], Blass [132], Freyd [373], Johnstone & Linton [543], Kock et al. [635].

D5.5 Orbitals and numerals

We observed at the start of the previous section that, in a topos without a natural number object, we cannot hope to have an object parametrizing all the 'finite' objects of the topos, for any reasonable notion of finiteness. Nevertheless, it does seem reasonable to expect that we can find some non-local notion of 'counting' in such toposes. The theory of numerals, introduced by P. Freyd [373], provides such a notion, which relates well to K-finiteness (and to cardinal-finiteness, when the topos does have a natural number object). In this account, we shall approach it via the theory of orbitals, which was developed more recently (but independently) by J. Bénabou and B. Loiseau [106].

The starting-point of the theory is an idea which we have already exploited in 5.1.1. As in Section D5.1, we shall write $\mathbb{T}$ for the algebraic theory freely generated by one nullary and one unary operation.

Definition 5.5.1 By an *orbital* in a topos $\mathcal{E}$, we mean a $\mathbb{T}$-algebra (A, a, t) which 'satisfies Peano's fifth postulate' internally, i.e. satisfies the sentence

$$(\forall w\colon PA)((w \text{ is } (a,t)\text{-closed}) \Rightarrow (w = \ulcorner A \urcorner)) .$$

In this terminology, then, 5.1.1 says that every $\mathbb{T}$-algebra contains a unique orbital (and that in an orbital the morphisms a and t are jointly epic). Hence we do not have to worry about interpreting the displayed sentence internally; if it is

externally true that a $\mathbb{T}$-algebra (A, a, t) has no proper (a, t)-closed subobjects, then the orbital which it contains must be the whole of A.

Lemma 5.5.2

(i) *Let (A, a, t) and (B, b, u) be orbitals in a topos $\mathcal{E}$. Then there exists at most one $\mathbb{T}$-algebra morphism $A \to B$, and if such a morphism exists it is an epimorphism (in $\mathcal{E}$).*

(ii) *If (A, a, t) is an orbital and $f\colon (A, a, t) \twoheadrightarrow (B, b, u)$ is an epimorphism (in $\mathcal{E}$) and a $\mathbb{T}$-algebra homomorphism, then (B, b, u) is an orbital.*

(iii) *If we write $(A, a, t) \geq (B, b, u)$ when there exists a morphism of $\mathbb{T}$-algebras $(A, a, t) \to (B, b, u)$, then the resulting preorder $\mathrm{Orb}(\mathcal{E})$ has finite joins and finite nonempty meets; and it has a greatest element iff $\mathcal{E}$ has a natural number object.*

(iv) *If we define an $\mathcal{E}$-indexed category $\mathbb{O}\mathsf{rb}(\mathcal{E})$ by $\mathrm{Orb}(\mathcal{E})^I = \mathrm{Orb}(\mathcal{E}/I)$, then this category has $\mathcal{E}$-indexed colimits; and it has $\mathcal{E}$-indexed limits if $\mathcal{E}$ has a natural number object.*

Proof (i) Given two $\mathbb{T}$-algebra morphisms $f, g\colon A \rightrightarrows B$, their equalizer is an (a, t)-closed subobject of A, and so must be the whole of A; similarly, the image of any homomorphism f must be a (b, u)-closed subobject of B.

(ii) Pullback along f maps (b, u)-closed subobjects to (a, t)-closed subobjects, and it preserves properness of subobjects.

(iii) The terminal object 1 of $\mathbb{T}$-**Mod**$(\mathcal{E})$ is clearly an orbital, and provides a least element of $\mathrm{Orb}(\mathcal{E})$. To form the join $(A, a, t) \cup (B, b, u)$ – we shall denote joins and meets in $\mathrm{Orb}(\mathcal{E})$ by $\cup$ and $\cap$ rather than $\vee$ and $\wedge$, for reasons which will become clear later – we observe that $(A \times B, (a, b), t \times u)$ is a $\mathbb{T}$-algebra, and so contains an orbital $A \cup B$ which admits morphisms to both A and B. Moreover, if (C, c, v) is another orbital which does this, then the image of the induced morphism $C \to A \times B$ is (an orbital, and hence) equal to $A \cup B$. So $A \cup B$ is a least upper bound for A and B in $\mathrm{Orb}(\mathcal{E})$. To form the meet of A and B, consider the pushout

$$\begin{tikzcd} A \cup B \arrow[r, two heads] \arrow[d, two heads] & A \arrow[d, two heads] \\ B \arrow[r, two heads] & A \cap B \end{tikzcd}$$

in $\mathcal{E}$. It is easily seen that $A \cap B$ can be given a unique $\mathbb{T}$-algebra structure such that the morphisms $A \twoheadrightarrow A \cap B$ and $B \twoheadrightarrow A \cap B$ are homomorphisms, and it is an orbital by (ii). But any pair of morphisms $(A \to C, B \to C)$ where C is an orbital must factor through this pushout, since the category of orbitals is a preorder; hence $A \cap B$ is a greatest lower bound for A and B. Finally, if N exists then it is initial in the category of $\mathbb{T}$-algebras and hence (since we have reversed

the arrows) a greatest element in $\mathrm{Orb}(\mathcal{E})$; conversely, a greatest element, if it exists, admits a unique homomorphism to each orbital and hence (by 5.1.1) to each $\mathbb{T}$-algebra, so it is a natural number object.

(iv) Since the property of being an orbital is defined by a sentence in the internal language, it is preserved by logical functors such as pullback functors; so the preorders $\mathrm{Orb}(\mathcal{E}/I)$ do indeed fit together to form an $\mathcal{E}$-indexed preorder. We may similarly observe that the process of cutting down a $\mathbb{T}$-algebra to its orbital, which provides a right adjoint to the inclusion $\mathrm{Orb}(\mathcal{E})^{\mathrm{op}} \to \mathbb{T}\text{-}\mathbf{Mod}(\mathcal{E})$, commutes with pullback functors; so $\mathbb{O}\mathrm{rb}(\mathcal{E})^{\mathrm{op}}$ is coreflective in the $\mathcal{E}$-indexed category of $\mathbb{T}$-algebras, and hence inherits $\mathcal{E}$-completeness (and $\mathcal{E}$-cocompleteness, if $\mathcal{E}$ has a natural number object) from the latter. □

Now suppose given an object A equipped with an endomorphism $t\colon A \to A$. We may form the object $C(A,t) \rightarrowtail PA$ of t-closed subobjects of A (i.e. those $A' \rightarrowtail A$ satisfying $A' \leq t^*(A')$). This subobject is clearly closed under arbitrary (internal) intersections in PA, and so the inclusion $C(A,t) \rightarrowtail PA$ has a left adjoint, giving rise to a closure operation $c\colon PA \to PA$. We shall write $O\colon A \to PA$ for the composite of c with the singleton map $\{\}\colon A \to PA$, and call it the *orbit map* associated with t. Clearly, for any $a\colon 1 \to A$, the subobject of A named by the composite Oa is an orbital (with distinguished element a).

We define a binary relation $\preceq$ on A by setting $x \preceq y$ iff $O(y) \leq O(x)$; i.e. $\preceq\ \rightarrowtail A \times A$ is the pullback along $O \times O$ of the opposite of the usual order-relation on PA. We also write $x \prec y$ for $tx \preceq y$. By the *cyclic part* $C(A)$ of A, we mean the subobject $\{x : A \mid x \prec x\}$; we say (A,t) is *acyclic* if its cyclic part consists entirely of fixed points of t, i.e. $C(A) \rightarrowtail A$ is the equalizer of t and 1_A.

Lemma 5.5.3 *With the above notation,*

(i) *$\preceq$ is a preorder on A.*

(ii) *If $f\colon (A,t) \to (B,u)$ is a morphism of objects-with-an-endomorphism, then f is an order-preserving map $(A,\preceq) \to (B,\preceq)$. In particular, t is an order-preserving map $(A,\preceq) \to (A,\preceq)$.*

(iii) *We have $(\forall x,y\colon A)((x \preceq y) \Leftrightarrow ((x \prec y) \vee (x = y)))$.*

(iv) *If (A,a,t) is an orbital, then it satisfies $(\forall x : A)(a \preceq x)$ and $(\forall x,y\colon A)((x \preceq y) \vee (y \preceq x))$.*

(v) *(A,t) is acyclic iff $\preceq$ is a partial order on A.*

Proof (i) It is clear that $\preceq$ inherits reflexivity and transitivity from the ordering on PA.

(ii) It is also clear that if f satisfies the hypothesis, then, for any $x\colon A$, $Of(x)$ is the image of $O(x)$ under f, since it is an orbital with distinguished element

$f(x)$. In other words, the diagram

$$
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "O"] & B \arrow[d, "O"] \\
PA \arrow[r, "\exists f"] & PB
\end{tikzcd}
$$

commutes. So the result follows from the fact that $\exists f$ is order-preserving. The second assertion follows from the fact that t commutes with itself.

(iii) It is again clear that $x \preceq y$ is equivalent to $y \in O(x)$; but since $O(x)$ is an orbital, this is in turn equivalent to $((x = y) \vee (\exists z)((z \in O(x)) \wedge (tz = y)))$ by 5.1.1. From the second half of this disjunction, we deduce $(tx \preceq y)$ (i.e. $(x \prec y)$) since t is order-preserving; so we have proved the left-to-right implication in (iii). But the right-to-left implication follows trivially from (i), since $x \preceq tx$.

(iv) The first assertion is clear since Oa is the whole of A. For the second, consider the subobject $A' \rightarrowtail A$ which is the interpretation of

$$\{x \colon A \mid (\forall y \colon A)((x \preceq y) \vee (y \preceq x))\};$$

we shall show that A' is (a, t)-closed, and hence the whole of A. Clearly a factors through A', by the result just established. Suppose $(x \in \ulcorner A' \urcorner)$; then for any y we have $((x \preceq y) \vee (y \preceq x))$. But by (iii) the first possibility is equivalent to $((x = y) \vee (tx \preceq y))$; and both $(x = y)$ and $(y \preceq x)$ imply $(y \preceq tx)$. So we have shown that $((x \in \ulcorner A' \urcorner) \Rightarrow (tx \in \ulcorner A' \urcorner))$.

(v) Suppose (A, t) is acyclic. Then from $((x \preceq y) \wedge (y \preceq x))$ we may deduce $((x = y) \vee ((x \prec y) \wedge (y \preceq x)))$, by (iii), and hence $((x = y) \vee (x \prec x))$, i.e. $((x = y) \vee (x \in \ulcorner C(A) \urcorner))$. But from $(x \in \ulcorner C(A) \urcorner)$ we deduce $(x = tx)$ and hence $(O(x) = \{x\})$ since $\{x\}$ is t-closed; so this possibility also leads to $(x = y)$. So we have shown that $\preceq$ is antisymmetric, and hence a partial order. Conversely, suppose $\preceq$ is antisymmetric; then from $(x \prec x)$ we deduce $((x \preceq tx) \wedge (tx \preceq x))$, and hence $(x = tx)$. So $C(A)$ is contained in the equalizer of t and 1_A; but the reverse inclusion always holds. □

We define a *numeral* to be an acyclic orbital. This is not Freyd's original definition; he took the partial order $\preceq$, as well as the morphisms a and t, as part of the primitive structure of a numeral. However, the following lemma shows that Freyd's definition is equivalent to the one just given.

Lemma 5.5.4 *Let (A, a, t) be an orbital, and suppose given a partial order $\leq$ on A such that t is order-preserving and a is the least element of A. Then $\leq$ coincides with $\preceq$; in particular, (A, a, t) is a numeral.*

Proof First we observe that we have $(\forall x \colon A)(x \leq tx)$, since the subobject of A which is the interpretation of $\{x \colon A \mid x \leq tx\}$ is (a, t)-closed by the hypotheses

on $\leq$. It is now clear that $(x \preceq y)$ implies $(x \leq y)$, since $\{y : A \mid x \leq y\}$ is a t-closed subobject of A containing x, and hence contains $O(x)$. For the converse, consider the subobject $A' \rightarrowtail A$ which is the interpretation of

$$\{x : A \mid (\forall y : A)((x \leq y) \Rightarrow (x \preceq y))\};$$

we shall show that A' is (a,t)-closed, and hence the whole of A. Clearly, a factors through A', since Oa is the whole of A; so suppose $(x \in \ulcorner A' \urcorner)$, and let $(tx \leq y)$. Since $(x \leq tx)$, we have $(x \leq y)$ and hence $(x \preceq y)$, so by 5.5.3(iii) we have $((x = y) \vee (x \prec y))$. But from the first possibility we deduce $(tx \leq x)$, whence $(x = tx)$ by antisymmetry of $\leq$, and so $(tx \preceq y)$; and the second possibility is exactly the statement that $(tx \preceq y)$. Thus we have shown that $((x \in \ulcorner A' \urcorner) \Rightarrow (tx \in \ulcorner A' \urcorner))$, i.e. A' is t-closed. □

Given an arbitrary orbital (A, a, t), we may consider the equivalence relation $\approx$ on A which is the intersection of $\preceq$ and its opposite. The quotient $\overline{A} = A/\approx$ clearly inherits a $\mathbb{T}$-algebra structure from A; and it is an orbital since we have an epimorphism $A \twoheadrightarrow \overline{A}$. Since it also carries a partial order (the image of $\preceq$) satisfying the conditions of 5.5.4, we deduce that it is a numeral. Moreover, any morphism from A to a numeral must factor through $A \twoheadrightarrow \overline{A}$, since it preserves $\preceq$ and hence $\approx$; so we have proved

Corollary 5.5.5 *The sub-preorder* $\mathrm{Num}(\mathcal{E})$ *of numerals in* $\mathcal{E}$ *is coreflective in* $\mathrm{Orb}(\mathcal{E})$. □

Moreover, the coreflection can be made into an $\mathcal{E}$-indexed coreflection, as in the proof of 5.5.2(iv). Thus the $\mathcal{E}$-indexed preorder $\mathsf{Num}(\mathcal{E})$ enjoys the same completeness properties as $\mathbb{O}\mathsf{rb}(\mathcal{E})$ (cf. 5.5.2(iv)).

Next, we prove a result which is the analogue for numerals of Freyd's characterization of natural number objects (5.1.2).

Proposition 5.5.6 *Let (A, a, t) be a $\mathbb{T}$-algebra in a topos $\mathcal{E}$, and $\leq$ a partial order on A such that t is order-preserving and a is the least element of A. Then (A, a, t) is a numeral iff it satisfies the following conditions:*

(i) *$(\forall x : A)(x \leq tx)$;*

(ii) *(a, t) is jointly epic;*

(iii) *$\leq$ is a total order;*

(iv) *$(\forall x, y : A)(((x \leq y) \wedge (y \leq tx)) \Rightarrow ((x = y) \vee (y = tx)))$;*

(v) *the coequalizer of t and 1_A is $A \to 1$.*

Proof First suppose (A, a, t) is a numeral. Then we have already seen that (i) holds in the proof of 5.5.4, (ii) holds by 5.1.1, and (iii) holds by 5.5.3(iv). For (iv), we observe that $(x \leq y)$ implies $((x = y) \vee (tx \leq y))$ by 5.5.3(iii) and 5.5.4, and that $((tx \leq y) \wedge (y \leq tx))$ implies $(y = tx)$ since $\leq$ is antisymmetric. Finally, for (v), suppose we have a morphism $f : A \to B$ satisfying $f = ft$. Then f is a

$\mathbb{T}$-algebra morphism $(A, a, t) \to (B, fa, 1_B)$, so its image is the orbit $O(fa) \rightarrowtail B$, which is clearly $\{fa\}$. Thus f factors through $A \to 1$ (and the factorization is unique since $A \to 1$ is split epic).

Conversely, suppose the five conditions hold. Let $A' \rightarrowtail A$ be the orbit named by Oa; we shall argue as in the proof of 5.1.2(ii) $\Rightarrow$ (iii) to show that A' is the whole of A. First of all, we claim that

$$(\forall x, y \colon A)(((x \leq y) \wedge (y \in \ulcorner A' \urcorner)) \Rightarrow (x \in \ulcorner A' \urcorner));$$

to prove this, let $A'' \rightarrowtail A'$ be the interpretation of

$$\{y \colon A' \mid (\forall x \colon A)((x \leq y) \Rightarrow (x \in \ulcorner A' \urcorner))\} .$$

We claim that A'' is (a, t)-closed, and hence the whole of A'. Clearly, a factors through A'' since $(x \leq a)$ implies $(x = a)$ by antisymmetry of $\leq$. If $(y \in A'')$ and $(x \leq ty)$, then we have either $(x \leq y)$ or $(y \leq x)$ by (iii); the first of these implies $(x \in \ulcorner A' \urcorner)$, and the second implies $((y = x) \vee (x = ty))$ by (iv), from which we again deduce $(x \in \ulcorner A' \urcorner)$. Next, we claim that

$$(\forall x \colon A)((x \in \ulcorner A' \urcorner) \Leftrightarrow (tx \in \ulcorner A' \urcorner));$$

the left-to-right implication is clear since A' is t-closed, so suppose $(tx \in \ulcorner A' \urcorner)$. From (ii), we deduce $((tx = a) \vee (\exists y : A')(tx = ty))$; but $(tx = a)$ implies $(a \leq x \leq tx = a)$ and hence $(x = a)$, so $(x \in \ulcorner A' \urcorner)$. Given y such that $((y \in \ulcorner A' \urcorner) \wedge (tx = ty))$, we have $((x \leq y) \vee (y \leq x))$ by (iii); the first possibility implies $(x \in A')$ by the argument already given, and the second implies $(y \leq x \leq tx = ty)$, whence $((y = x) \vee (x = ty))$ by (iv), and either of these possibilities implies $(x \in \ulcorner A' \urcorner)$.

Thus we have shown that A' is S-closed, where $S \rightarrowtail A \times A$ is the symmetrization of the graph of t (i.e. the union of the graph and its opposite). Arguing exactly as in the proof of 5.1.2(ii) $\Rightarrow$ (iii), we now deduce that it is T-closed, where T is the kernel-pair of the coequalizer of $S \rightrightarrows A$. But T is the whole of $A \times A$ by (v); so, since A' contains the element a, it must be the whole of A. □

Corollary 5.5.7 *Any bicartesian functor between toposes preserves numerals.*

Proof Such a functor preserves the validity of the conditions of 5.5.6 (note that the first four of them are expressible as coherent sequents, and are thus preserved by arbitrary coherent functors). □

Let A be an object of a topos equipped with a reflexive relation $R \colon A \nrightarrow A$. Then the composite of R with the universal relation $\in_A \colon PA \nrightarrow A$ (cf. A3.4.5) defines a morphism $r \colon PA \to PA$, which corresponds to the operation of composing an arbitrary relation $B \nrightarrow A$ with R. It is clear that r is order-preserving (for the usual ordering on PA) and inflationary (i.e. it satisfies $(\forall w \colon PA)(w \leq rw)$).

So if we fix any particular subobject $A' \rightarrowtail A$ and cut down the $\mathbb{T}$-algebra $(PA, \ulcorner A' \urcorner, r)$ to an orbital, the latter will be a numeral by 5.5.4. Let us write B for this numeral, and consider the subobject $A'' \rightarrowtail A$ which is the internal union of the B-indexed family of subobjects named by $B \rightarrowtail PA$. We claim that A'' is the R-closure of A', that is the smallest subobject satisfying $A' \leq A''$ and

$$(\forall x, y \colon A)(((x \in \ulcorner A'' \urcorner) \wedge (\langle x, y \rangle \in \ulcorner R \urcorner)) \Rightarrow (y \in \ulcorner A'' \urcorner)) .$$

For A'' clearly has this closure property (since B is r-closed), and contains A' since $\ulcorner A' \urcorner$ factors through B. Conversely, if $C \rightarrowtail A$ is R-closed in the above sense, then $\ulcorner C \urcorner \colon 1 \to PA$ is a fixed point of r; hence if we also have $A' \leq C$ then $\{w \colon PA \mid \ulcorner A' \urcorner \leq w \leq \ulcorner C \urcorner\}$ is a sub-$\mathbb{T}$-algebra of $(PA, \ulcorner A' \urcorner, r)$ and so contains B, i.e. $(\forall w \colon B)(w \leq \ulcorner C \urcorner)$.

We may now deduce, for toposes not necessarily having natural number objects, the analogue of a result which we proved for toposes with natural number objects back in A2.5.7:

Proposition 5.5.8 *Let $T \colon \mathcal{E} \to \mathcal{F}$ be a coherent functor between toposes. Then the following conditions are equivalent:*

(i) *T is cocartesian (equivalently, preserves coequalizers);*
(ii) *T preserves R-closures of subobjects, for any reflexive relation R;*
(iii) *T preserves transitive closures of reflexive relations.*

Proof (i) $\Rightarrow$ (ii): Let $R \rightarrowtail A \times A$ be a reflexive relation in $\mathcal{E}$, $A' \rightarrowtail A$ an arbitrary subobject and $A'' \rightarrowtail A$ its R-closure. Since a coherent functor preserves composition of relations, it sends A'' to a TR-closed subobject of TA containing TA'; we have to show that it is no larger than the TR-closure $C \rightarrowtail TA$ of TA'. But we constructed A'' above as the image of $v \colon S \to A$, where $(u, v) \colon S \rightarrowtail B \times A$ is the relation corresponding to $B \rightarrowtail PA$; so TA'' is the image of $Tv \colon TS \to TA$. Now consider

$$\{y \colon TB \mid \{z \colon TA \mid \langle y, z \rangle \in \ulcorner TS \urcorner\} \leq \ulcorner C \urcorner\};$$

it is not hard to verify that this is a $(T\ulcorner A' \urcorner, Tr)$-closed subobject of TB, but TB is a numeral by 5.5.7 and so it is the whole of TB. Hence $TA'' \leq C$, as required.

(ii) $\Rightarrow$ (iii): Given a reflexive relation $R \rightarrowtail A \times A$, let $B = A \times A$ and let $\widetilde{R} \rightarrowtail B \times B$ be the relation defined by the term

$$\{\langle x, y, z, t \rangle \colon A^4 \mid ((x = z) \wedge (\langle y, t \rangle \in \ulcorner R \urcorner))\} .$$

If we take $B' \rightarrowtail B$ to be the diagonal subobject $A \rightarrowtail A \times A$, then it is clear that the $\widetilde{R}$-closure of B' is just the transitive closure of R.

(iii) $\Rightarrow$ (i): Any coherent functor between toposes preserves coequalizers of equivalence relations; and it preserves the construction of the reflexive and symmetric relation generated by a parallel pair of morphisms in $\mathcal{E}$, since the latter

may be constructed using finite limits, finite coproducts and images. Hence if (iii) holds, T preserves the construction of the equivalence relation generated by a parallel pair. □

So far in this section, we have not had much to say about finiteness as such. We now make good this deficit.

Lemma 5.5.9 *Let (A, a, t) be an orbital, and let $C(A)$ denote the cyclic part of A.*

(i) *(A, a, t) is a natural number object iff $C(A) \cong 0$.*

(ii) *A is K-finite iff $C(A)$ is well-supported.*

(iii) *The equalizer of t and 1_A is a subterminal object.*

(iv) *(A, a, t) is a K-finite numeral iff $C(A) \cong 1$.*

Proof (i) The left-to-right implication is clear, since the preorder $\preceq$ on N must coincide with the usual partial order by 5.5.4. Conversely, suppose $C(A) \cong 0$. Then the image of t is disjoint from a, since $(tx = a)$ would imply $(tx \preceq x)$, i.e. $(x \in \ulcorner C(A) \urcorner)$. So by 5.1.2 it suffices to prove that t is monic. Suppose $(tx = ty)$; by 5.5.3(iv) we have either $(x \preceq y)$ or $(y \preceq x)$. If the former holds, then we have $(x \preceq y \preceq ty = tx)$; but (A, t) is clearly acyclic, and so by 5.5.6(iv) we have $((x = y) \vee (y = tx))$. But $(y = tx)$ would imply $(y \in \ulcorner C(A) \urcorner)$, which is impossible, so we have $(x = y)$. The argument if $(y \preceq x)$ is similar.

(ii) First suppose A is K-finite. We shall say that a subobject $A' \rightarrowtail A$ is bounded if $(\exists y : A)(\forall x : A')(x \preceq y)$; clearly, we may construct a subobject $L \rightarrowtail PA$ indexing the bounded subobjects of A. This subobject clearly contains $\ulcorner 0 \rightarrowtail A \urcorner$ and singletons; and it is closed under binary unions by 5.5.3(iv). So $K(A) \leq L$ and hence A is bounded, i.e. $(\exists y \colon A)(\forall x \colon A)(x \preceq y)$. Substituting ty for x in this formula, we deduce that any such y must be cyclic.

For the converse, we first assume that (A, t) is acyclic. Suppose $(y \in \ulcorner C(A) \urcorner)$; then $\{x : A \mid (x \preceq y)\}$ defines an (a, t)-closed subobject of A (since $(x \preceq y)$ implies $(tx \preceq ty = y)$) which must be the whole of A, i.e. y must be an upper bound for A. Now consider the subobject $A' \rightarrowtail A$ defined by

$$\{y \colon A \mid \{x \colon A \mid (x \preceq y)\} \text{ is } K\text{-finite}\}\ ;$$

we shall show that A' is (a, t)-closed, and so any bounded subobject of A (in particular, A itself) is K-finite. Clearly a factors through A' since $\{x \colon A \mid (x \preceq a)\} = \{a\}$; and from 5.5.6(iii) and (iv) we obtain

$$\{x \colon A \mid (x \preceq ty)\} = \{x \colon A \mid (x \preceq y)\} \cup \{ty\},$$

whence $(\forall y \colon A)((y \in \ulcorner A' \urcorner) \Rightarrow (ty \in \ulcorner A' \urcorner))$.

Unfortunately, if (A, t) contains nontrivial cycles, the above argument will not work, since the sets $\{x : A \mid x \preceq y\}$ need not be K-finite in this case. For

such an A, we consider the morphism $\hat{t}\colon K(A) \to K(A)$ given by

$$K(A) \xrightarrow{(1_{K(A)}, K(t))} K(A) \times K(A) \xrightarrow{\cup} K(A),$$

and write $\widehat{O}\colon A \to P(K(A))$ for the composite of the singleton map $A \to K(A)$ and the orbit map for $\hat{t}$. We claim first that we have

$$(\forall x\colon A)(\forall w\colon K(A))((w \in \widehat{O}(x)) \Rightarrow (\exists y\colon A)(\hat{t}w = w \cup \{ty\}))\ .$$

We prove this, for a given x, by induction over w in $\widehat{O}(x)$: clearly, for $w = \{x\}$ we may take $y = x$, and if $w = \hat{t}w'$ and y' satisfies the condition for w', then ty' satisfies it for y. Similarly, we may prove

$$(\forall x, y\colon A)(\forall w\colon K(A))(((w \in \widehat{O}(x)) \wedge (y \in w)) \Rightarrow (O(x) = w \cup O(y)))\ .$$

The key idea is now to consider

$$\{w\colon K(A) \mid ((w \in \widehat{O}(x)) \wedge (\forall y\colon A)(((y \in w) \wedge (ty = x)) \Rightarrow (w = O(x))))\}$$

where x is a free variable of type A. Once again, we may show that this set contains $\{x\}$ and is $\hat{t}$-closed, so it is the whole of $\widehat{O}(x)$. We deduce that

$$((\exists y\colon A)((y \in O(x)) \wedge (ty = x)) \Rightarrow (O(x) \text{ is } K\text{-finite}));$$

but we know that $(\exists x\colon A)(tx \preceq x)$, and this particular x must equal ty for some $y \in O(x)$. Now since $Oa = \ulcorner A \urcorner$, we have

$$(\exists w\colon K(A))((w \in \widehat{O}(a)) \wedge (x \in w))$$

for this particular x, but we then have $\ulcorner A \urcorner = w \cup O(x)$. So A is the union of two K-finite subobjects, and hence K-finite.

(iii) Suppose $(x = tx)$; then, as in the proof of (ii), we deduce that x is an upper bound for A, i.e. $(\forall y\colon A)(x \in O(y))$. But if we also have $(y = ty)$ then $O(y) = \{y\}$; so any two fixed points of t must be equal.

(iv) If A is a K-finite numeral, then $C(A)$ is well-supported by (ii) and subterminal by (iii); so it is isomorphic to 1. Conversely, if $C(A) \cong 1$, then A is K-finite by (ii). Also, it is clear that if $(x \in \ulcorner C(A) \urcorner)$ then we also have $(tx \in \ulcorner C(A) \urcorner)$; so $C(A) \cong 1$ implies that any cyclic element x satisfies $(x = tx)$. □

Thus, in a 2-valued topos, every orbital is either K-finite or a natural number object. In the opposite direction, it is clear that every finite successor cardinal can be given the structure of a numeral: we can 'truncate' the successor morphism on N to yield a morphism $[sp] \to [sp]$ with the required properties. However, not every quotient of such a cardinal can be given a numeral structure or even

an orbital structure. For example, let $\mathcal{C}$ be the finite category with four objects a, b, c, d and three non-identity morphisms $a \to b$, $a \to c$, $a \to d$, and let $A \in$ ob $[\mathcal{C}, \mathbf{Set}]$ be the functor defined by $A(a) = \{0, 1, 2\}$, $A(b) = A(c) = A(d) = \{0, 1\}$ with $A(a \to b)$, $A(a \to c)$ and $A(a \to d)$ being surjections which identify the pairs $(1, 2)$, $(0, 2)$ and $(0, 1)$ respectively. Then it is easily seen that there is no way of defining a transitive 'successor map' on $A(a)$ which is compatible with all three transition maps. (If we delete one of the three 'legs' of this diagram, we obtain a K-finite object which admits an orbital structure but no numeral structure.)

But the class of K-finite numerals does contain more than just the finite cardinals:

Lemma 5.5.10

(i) *If $\mathcal{E}$ has a natural number object, then* $\mathsf{Num}(\mathcal{E})$ *is equivalent to the opposite of (the standard indexing of) the internal poset of s-closed subobjects of N.*

(ii) *Even if $\mathcal{E}$ does not have a natural number object,* $\mathsf{Num}(\mathcal{E})^{\mathrm{op}}$ *is cartesian closed.*

Proof (i) For any numeral (A, a, t), we have a unique $\mathbb{T}$-algebra morphism $g \colon N \twoheadrightarrow A$, and A is characterized up to isomorphism by the set $\{n \colon N \mid g(n) = g(sn)\}$. Conversely, given an s-closed subobject $B \rightarrowtail N$, the pushout

$$\begin{tikzcd} B \arrow[r, tail] \arrow[d, two heads] & N \arrow[d, two heads, "g"] \\ \sigma B \arrow[r, tail] & A \end{tikzcd}$$

inherits a numeral structure from N; and since this square is also a pullback by A2.4.3, we can recover B as $\{n : N \mid g(n) = g(sn)\}$. Thus we have an order-reversing bijection between isomorphism classes of numerals in $\mathcal{E}$ and s-closed subobjects of N; and the same argument works in $\mathcal{E}/I$ for any I, so we have an equivalence of $\mathcal{E}$-indexed categories.

(ii) It suffices to prove that, given any numeral (A, a, t), the preorder $\{B \in \text{ob Num}(\mathcal{E}) \mid B \leq A\}^{\mathrm{op}}$ is cartesian closed. But, as in (i), we may identify these numerals up to isomorphism with subobjects of A which are t-closed (equivalently, which are upwards closed for the relation $\preceq$) and contain the subobject $\{x \colon A \mid (x = tx)\}$. These clearly form a sublattice of $\mathrm{Sub}(A)$; they do not form a sub-Heyting-algebra, but they do have an implication defined by

$$(B \Rightarrow C) = \{x \colon A \mid (\forall y)(((x \preceq y) \wedge (y \in \ulcorner B \urcorner)) \Rightarrow (y \in \ulcorner C \urcorner))\} .$$

$\square$

In the situation of 5.5.10(i), it is clear that the K-finite numerals correspond to the inhabited s-closed subobjects of N, and the inhabited finite cardinals correspond to the s-closed subobjects with least elements – equivalently, to the complemented ones. Thus, for example, whereas we saw in 5.2.8(*b*) that cardinals in a spatial topos $\mathbf{Sh}(X)$ correspond to continuous (= locally constant) functions $f\colon X \to \mathbb{N}$, the finite numerals in this topos correspond to (everywhere-nonzero) upper semicontinuous functions, i.e. those for which $\{x \in X \mid f(x) \leq n\}$ is open for each n. (Explicitly, a numeral A corresponds to the function whose value at $x \in X$ is the number of elements in the stalk A_x.)

There is another important connection between numerals and K-finiteness. Given an arbitrary object A of a topos, consider the relation $R \rightarrowtail PA \times PA$ defined by the term

$$\{z\colon PA \times PA \mid ((\mathsf{fst}(z) = \mathsf{snd}(z)) \vee (\exists x\colon A)(\mathsf{snd}(z) = \mathsf{fst}(z) \cup \{x\}))\} .$$

Clearly, R is reflexive, and so we can use the ideas described before 5.5.8 to construct the R-closure of $\ulcorner 0 \urcorner\colon 1 \rightarrowtail PA$; but, by 5.4.2, the latter is precisely $K(A)$. Thus we have a numeral $(n(A), o, s)$ (which we shall call the numeral associated with A) and a relation $\phi\colon n(A) ⇸ K(A)$ satisfying the coherent sequents $(\phi(o, w) \dashv\vdash_w (w = \ulcorner 0 \urcorner))$ and

$$(\phi(sn, w) \dashv\vdash_{n,w} (\phi(n, w) \vee (\exists w'\colon K(A))(\exists x\colon A)(\phi(n, w') \wedge (w = (w' \cup \{x\})))))$$

(where n and w are variables of types $n(A)$ and $K(A)$ respectively). The idea is that $n(A)$ 'counts the number of elements required to cover each finite subobject of A'.

Proposition 5.5.11 *Let A be a K-finite object. Then*

(i) *$n(A)$ is K-finite;*

(ii) *$n(A)$ is characterized up to isomorphism in* $\mathrm{Num}(\mathcal{E})$ *by the sequents above together with*

$$(\phi(n, \ulcorner A \urcorner) \dashv\vdash_n (n = sn));$$

(iii) *if A underlies a numeral (A, a, t), then $n(A)$ is isomorphic to the numeral $(A, a, t)^+$ with underlying object $1 \amalg A$ and structure maps $\nu_1 : 1 \to 1 \amalg A$ and $\nu_2(a, t)\colon 1 \amalg A \to 1 \amalg A$.*

Proof (i) We first establish the validity of

$$(\forall n\colon n(A))(\forall w, w'\colon K(A))((\phi(n, w) \wedge (w' \leq w)) \Rightarrow \phi(n, w'))$$

by showing that the set of n for which the condition holds is a (o, s)-closed subobject of $n(A)$. Clearly, the condition holds for o since $\ulcorner 0 \urcorner$ is the least element of $K(A)$. Suppose it holds for n; let $B \rightarrowtail A$ be a K-finite subobject for which we have $\phi(sn, \ulcorner B \urcorner)$ and $B' \rightarrowtail A$ a K-finite subobject contained in B. Then we have either $\phi(n, \ulcorner B \urcorner)$ or $B = B'' \cup \{a\}$ for some $a\colon 1 \to A$ and some K-finite

B'' such that $\phi(n, \ulcorner B'' \urcorner)$. In the first case, we clearly have $\phi(n, \ulcorner B' \urcorner)$ and hence $\phi(sn, \ulcorner B' \urcorner)$. In the second case, things are more complicated: because the union $B'' \cup \{a\}$ is not necessarily disjoint, the intersections $B' \cap B''$ and $B' \cap \{a\}$ need not be K-finite. However, we can prove

$$(\forall w \colon K(B))((w \leq \ulcorner B'' \urcorner) \vee (\exists w' \colon K(B))((w' \leq \ulcorner B'' \urcorner) \wedge (w = w' \vee \ulcorner \{a\} \urcorner)))$$

by showing that the object of K-finite subobjects of B satisfying this condition is a join-semilattice and contains singletons; so we have either $B' \leq B''$ or there exists a K-finite $C \leq B''$ such that $B' = C \cup \{a\}$. In either case, we deduce $\phi(sn, \ulcorner B' \urcorner)$, as required.

Now since A is K-finite, there exists n such that $\phi(n, \ulcorner A \urcorner)$ holds. But this n satisfies $(\forall w \colon K(A))\phi(n, w)$, from which it follows that $n = sn$; so $n(A)$ has an s-fixed point, and is thus K-finite by 5.5.9(ii).

(ii) If we have $\phi(n, \ulcorner A \urcorner)$, then by (i) we have $n = sn$. Conversely, if $n = sn$, then n is the $\preceq$-largest element of $n(A)$, and so $\phi(n, \ulcorner A \urcorner)$ follows from the fact that the mapping $n \mapsto \{w \colon K(A) \mid \phi(n, w)\}$ is order-preserving.

Conversely, if (B, b, t) is any numeral equipped with a relation $\psi \colon B \nrightarrow K(A)$ satisfying the sequents before the statement of the proposition, then the name $\ulcorner \psi \urcorner \colon B \to PPA$ of the composite relation $B \nrightarrow K(A) \rightarrowtail PA$ is easily seen to be a morphism of $\mathbb{T}$-algebras, where PPA is given the $\mathbb{T}$-algebra structure $(\{\ulcorner 0 \urcorner\}, r)$ used in the definition of $n(A)$. So its image is precisely $n(A)$. The additional condition in the statement of (ii) says precisely that, as a morphism $B \to n(A)$, $\ulcorner \psi \urcorner$ reflects the cyclic element; but since we have

$$(\forall y, z \colon B)((y = z) \vee (y \prec z) \vee (z \prec y))$$

by 5.5.3(iii) and (iv), this easily implies that $\ulcorner \psi \urcorner$ is injective and hence an isomorphism.

(iii) Let $\downarrow \colon A \to K(A)$ be the morphism defined by the term

$$\{y \colon A \mid y \preceq x\}$$

where x is a free variable of type A (note that the sets of this form are indeed K-finite, by the proof of 5.5.9(ii)). Now we claim that

$$(\forall x \colon A)(\exists n \colon n(A))(\phi(n, \downarrow x) \wedge (\forall m \colon n(A))(\phi(m, \downarrow x) \Rightarrow (n \preceq m))) \;;$$

we prove this by showing that the subobject of those x for which the condition holds is (a, t)-closed. The condition holds for $x = a$, since we have $\downarrow a = \{a\}$, and hence $\phi(so, \downarrow a)$ (but not $\phi(o, \downarrow a)$). Suppose $x = ty$ and that the condition holds for y; let n be the $\preceq$-least element of $n(A)$ satisfying $\phi(n, \downarrow y)$. Since $\downarrow x = \downarrow y \cup \{x\}$, it is clear that we have $\phi(sn, \downarrow x)$; also, any m for which $\phi(m, \downarrow x)$ holds also satisfies $\phi(m, \downarrow y)$ and hence $n \preceq m$. And if $\phi(n, \downarrow x)$ holds, then we must have $\downarrow x = \downarrow y$, whence $ty = y$, so $\downarrow y$ is the whole of A and hence $sn = n$. Thus we see that sn is the $\preceq$-least element of $n(A)$ satisfying $\phi(sn, \downarrow x)$.

Since the least n associated with x by the above formula is clearly unique, we have thus defined a morphism $f: A \to n(A)$, which we may extend to a morphism $\overline{f}: 1 \amalg A \to n(A)$ by sending the extra element to o. From the construction, it is clear that $\overline{f}$ is monic, and that it is a morphism of $\mathbb{T}$-algebras, where $1 \amalg A$ is given the structure maps in the statement of the proposition; so by 5.5.2(i) it is an isomorphism. $\square$

Following on from 5.5.11(i), we define an object A of a topos to be *R-finite* if $n(A)$ is K-finite. (R stands for 'Russell'; intuitively, A is R-finite if there exists a finite bound for the sizes of its K-finite subobjects.) We note some of the basic properties of this finiteness notion:

Lemma 5.5.12

(i) *An object A is R-finite iff there exists a K-finite numeral (B, b, t) and a relation $\psi: B \looparrowright K(A)$ satisfying the sequents $(\psi(b, w) \dashv\vdash_w (w = \ulcorner 0 \urcorner))$ and*

$$(\psi(ty, w) \dashv\vdash_{y,w} (\psi(y, w) \vee (\exists w' : K(A))(\exists x : A)(\phi(y, w') \wedge (w = (w' \cup \{x\}))))) .$$

(ii) *Any subobject of an R-finite object is R-finite.*

(iii) *Any $\tilde{K}$-finite object (cf. 5.4.22) is R-finite.*

(iv) *A numeral is R-finite iff it is K-finite.*

(v) *Bicartesian functors between toposes preserve R-finiteness.*

Proof (i) One direction is immediate by taking $B = n(A)$. Conversely, if ψ satisfies the conditions, then as in the proof of 5.5.11(ii) we see that $\ulcorner \psi \urcorner : B \to PPA$ induces a morphism of numerals $B \to n(A)$, so $n(A)$ inherits K-finiteness from B.

(ii) If A' is a subobject of A, we may identify $K(A')$ with the sub-semilattice of $K(A)$ defined by the term $\{w : K(A) \mid (\forall x : A)((x \in w) \Rightarrow (x \in \ulcorner A' \urcorner))\}$. It is then easy to see that the relation $\phi: n(A) \looparrowright K(A)$ restricts to a relation $n(A) \looparrowright K(A')$ having the properties in (i); so A' is R-finite.

(iii) By (ii) and 5.5.11(i), every subobject of a K-finite object is R-finite. But R-finiteness is clearly a local property, so the result follows from 5.4.22.

(iv) is immediate from 5.5.11(iii) (whose proof did not in fact make use of the K-finiteness of A).

(v) follows easily from (i), since bicartesian functors preserve K-finite numerals by 5.5.7 and 5.4.12. $\square$

Example 5.5.13 The converse of 5.5.12(iii) is false. Let $W \rightarrowtail \Omega$ be the Higgs object, as constructed in Section D4.5: it follows easily from the first displayed formula in the proof of 4.5.12 that W satisfies

$$(\forall x, y, z : W)((x = y) \vee (y = z) \vee (x = z)),$$

from which it follows that any inhabited K-finite subobject of W is a quotient of 2, and hence that $n(W) \leq 3$ in $\mathrm{Num}(\mathcal{E})$; so W is R-finite in any topos. But

we identified the Higgs object of the functor category $[\mathbf{N}, \mathbf{Set}]$ (where $\mathbf{N}$ is the ordered set of natural numbers) at the end of 5.4.25, and we observed that it is not $\tilde{K}$-finite.

Remark 5.5.14 . We may now develop an 'arithmetic of numerals' which parallels and extends the arithmetic of finite cardinals in a topos with natural number object. We have already defined a successor operation $(A, a, t) \mapsto (A, a, t)^+$ on numerals; the corresponding 'predecessor operation' $(-)^-$ sends (A, a, t) to the image of t, equipped with the structure maps ta and the restriction of t to its image. Because numerals are necessarily well-supported, we have a slightly awkward 'shift' between cardinals and numerals: we have to regard 1 as the numeral version of zero (it is, of course, $n(0)$), and $1 \amalg 1 = n(1)$ as the numeral version of one. We define the sum $(A, a, t) \oplus (B, b, u)$ to be $n(A \vee B)^-$, where $A \vee B$ is the pushout

$$\begin{tikzcd} 1 \arrow[r, "b"] \arrow[d, "a"] & B \arrow[d] \\ A \arrow[r] & A \vee B \end{tikzcd} \quad ;$$

note that $A \vee B$ is K-finite if A and B are, since it is a quotient of $A \amalg B$. The morphisms $1_A \times b \colon A \cong A \times 1 \to A \times B$ and $a \times 1_B \colon B \cong 1 \times B \to A \times B$ induce a monomorphism $A \vee B \rightarrowtail A \times B$, and we may thus define the product $(A, a, t) \otimes (B, b, u)$ to be $n(A \wedge B)^-$, where $A \wedge B$ is the pushout

$$\begin{tikzcd} A \vee B \arrow[r, tail] \arrow[d] & A \times B \arrow[d] \\ 1 \arrow[r] & A \wedge B \end{tikzcd}$$

(again, this is K-finite if A and B are, since it is a quotient of $A \times B$). We can even define truncated subtraction: given numerals A and B; the least numeral C satisfying $C \oplus B \geq A$ is given by $n(D)^-$, where D is the 'kernel' of the $\mathbb{T}$-algebra epimorphism $A \twoheadrightarrow A \cap B$, i.e. the s-closed subobject of A which corresponds to $A \cap B$ as in the proof of 5.5.10(ii). Note that D need not be K-finite if A is, since it need not be complemented in A; however, $n(D)$ is still K-finite, since D is certainly a subobject of a K-finite object.

It is not hard to verify that the operations on numerals thus defined satisfy (most of!) the usual laws of arithmetic, and that if $\mathcal{E}$ does have a natural number object then the mapping $p \mapsto [sp] = n([p])$ is a 'homomorphism' $\mathcal{E}(1, N) \to \mathrm{Num}(\mathcal{E})$. However, some familiar identities involving $\dot{-}$ fail to hold for numerals, because of the non-Boolean nature of their arithmetic: for example, on the subset

$\{A \in \mathrm{Num}(\mathcal{E}) \mid A \leq n(1)\}$, the operations $\oplus$ and
with the lattice operations $\cup$ and $\cap$, and hence $\dot{-}$
implication on this preorder (which is equivalent
of 5.5.10). So, for example, $n(1) \dot{-} (n(1) \dot{-} A)$ ne
$A \leq n(1)$. We shall leave the further exploration
to the reader.

Suggestions for further reading: Bénabou &

$\otimes$ are readily seen to coincide
coincides with the co-Heyting
to $\mathrm{Sub}_{\mathcal{E}}(1)^{\mathrm{op}}$, by the argument
ed not be isomorphic to A if
of the arithmetic of numerals

Loiseau [106], Freyd [373].

BIBLIOGRAPHY

[1] O. Acuña-Ortega, Some results in the theory of lattices in an arbitrary topos (Spanish), *Cienc. Tecn.* 2 (1978), 3–6; MR 82c:18004.

[2] O. Acuña-Ortega, A note on finiteness in an elementary Boolean topos (Spanish), *Cienc. Tecn.* 2 (1978), 7–11; MR 82c:18005.

[3] O. Acuña-Ortega, *Tertium non datur* in a topos of finite sets (Spanish), *Cienc. Tecn.* 3 (1979), 21–26; MR 84a:18005a.

[4] O. Acuña-Ortega, On a characterization of toposes of finite cardinals (Spanish), *Cienc. Tecn.* 4 (1980), 3–6; MR 84a:18005b.

[5] O. Acuña-Ortega, A characterization of k-finite decidable objects, *Cienc. Tecn.* 6 (1982), 3–10; MR 84f:18007.

[6] O. Acuña-Ortega, Internally projective objects and equivalence relations, *Cienc. Tecn.* 6 (1982), 11–15; MR 84f:18008.

[7] O. Acuña-Ortega, A finiteness theorem of Tarski in an arbitrary topos E, *Cienc. Tecn.* 7 (1983), 3–14; MR 85i:03194.

[8] O. Acuña-Ortega, Automorphisms of finite cardinals, *Cienc. Tecn.* 7 (1983), 15–28; MR 85g:03094.

[9] O. Acuña-Ortega, Finite cardinals in an arbitrary topos (Spanish), *Mat. Costarricense* 1 (1984), 7–28; MR 85d:03126.

[10] O. Acuña-Ortega, An exact coexact characterization of the finite cardinals, *J. Pure Appl. Alg.* 43 (1986), 211–233; MR 88j:18003.

[11] O. Acuña-Ortega, Finite objects and extensional relations, *Cahiers Top. Géom. Diff. Catégoriques* 28 (1987), 173–181; MR 88k:18007.

[12] O. Acuña-Ortega, Finite objects and automorphisms, *Commun. Alg.* 20 (1992), 3459–3478; MR 93k:18008.

[13] O. Acuña-Ortega, Definability and finite cardinals, *FernUniversität Hagen, Seminarberichte aus dem Fachbereich Mathematik* 63 (1998), 529–544.

[14] O. Acuña-Ortega and F. E. J. Linton, Finiteness and decidability, I, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 80–100; MR 81f:18016a.

[15] J. Adámek, On quasitopos hulls of topological categories, *Rend. Circ. Mat. Palermo Suppl.* (II) 24 (1990), 21–33; MR 92g:54010.

[16] J. Adámek, How to sketch quasivarieties, *J. Algebra* 185 (1996), 643–659; MR 98k:03041.

[17] J. Adámek and H. Herrlich, A characterization of concrete quasitopoi by injectivity, *J. Pure Appl. Alg.* 68 (1990), 1–9; MR 91m:18006.

[18] J. Adámek, H. Herrlich and J. Rosický, Essentially equational categories, *Cahiers Top. Géom. Diff. Catégoriques* 29 (1988), 175–192; MR 90b:18008.

[19] J. Adámek, P. T. Johnstone, J. A. Makowsky and J. Rosický, Finitary sketches, *J. Symbolic Logic* 62 (1997), 699–707; MR 99e:03044.

[20] J. Adámek, V. Koubek and V. Trnková, How large are left exact functors?, *Theory Appl. Categ.* 8 (2001), 377–390.

[21] J. Adámek and J. Reiterman, The quasitopos hull of the category of uniform spaces, *Topology Appl.* 27 (1987), 97–104; MR 88j:54014; corrigendum *ibid.* 48 (1992), 179–184; MR 93m:54023.

[22] J. Adámek, J. Reiterman and F. Schwarz, On universally topological hulls and quasitopos hulls, *Houston J. Math.* 17 (1991), 375–383; MR 93a:18002.

[23] J. Adámek and J. Rosický, *Locally presentable and accessible categories*, L.M.S. Lecture Notes Series no. 189 (Cambridge Univ. Press, 1994); MR 95j:18001.

[24] J. Adámek and J. Rosický, Finitary sketches and finitely accessible categories, *Math. Struct. Comput. Sci.* 5 (1995), 315–322; MR 97b:03086.

[25] J. Adámek and J. Rosický, On geometric and finitary sketches, *Appl. Categ. Struct.* 4 (1996), 227–240; MR 98j:18004.

[26] J. Adámek and J. Rosický, On sifted colimits and generalized varieties, *Theory Appl. Categ.* 8 (2001), 33–53; MR 01m:18001.

[27] J. Adamek and W. Tholen, Total categories with generators, *J. Algebra* 133 (1990), 63–78; MR 91h:18004.

[28] J. Adámek and W. Tholen, Totality of product completions, *Comment. Math. Univ. Carolinae* 41 (2000), 9–24; MR 01b:18002.

[29] M. Adelman, Points in topoi of sheaves over distributive lattices, *Bull. Austral. Math. Soc.* 20 (1979), 273–279; MR 81e:18007.

[30] M. Adelman and A. R. Blass, Exact functors, local connectedness and measurable cardinals, *Rend. Sem. Mat. Fis. Milano* 54 (1984), 9–28; MR 88j:18002.

[31] M. Adelman and J. V. Corbett, A sheaf model for intuitionistic quantum mechanics, *Appl. Categ. Struct.* 3 (1995), 79–104; MR 96a:81010.

[32] M. Adelman and P. T. Johnstone, Serre classes for toposes, *Bull. Austral. Math. Soc.* 25 (1982), 103–115; MR 84a:18018.

[33] P. Ageron, The logic of structures, *J. Pure Appl. Alg.* 79 (1992), 15–34; MR 93e:03083.

[34] C. Anghel and P. Lecouturier, Généralisation d'un résultat sur le triple de la réunion, *Ann. Fac. Sci. de Kinshasa (Zaire), Sect. Math.-Phys.* 1 (1975), 65–94; MR 53/582.

[35] M. Artin, *Grothendieck topologies*, Seminar notes, Harvard University (1962).

[36] M. Artin, A. Grothendieck and J. L. Verdier, *Théorie des topos et cohomologie étale des schémas*, Séminaire de Géométrie Algébrique du Bois-Marie, année 1963–64; second edition published as Lecture Notes in Math. vols. 269, 270 and 305 (Springer–Verlag, 1972); MR 50/7130–2.

[37] S. Awodey, Structures in mathematics and logic: a categorical perspective, *Philos. Math.* (3) 4 (1996), 209–237; MR 99c:03106.

[38] S. Awodey, Sheaf representation for topoi, *J. Pure Appl. Alg.* 145 (2000), 107–121; MR 01a:18003.

[39] S. Awodey and C. Butz, Topological completeness for higher-order logic, *J. Symbolic Logic* 65 (2000), 1168–1182.

[40] A. Bagchi, De Morgan's law and related identities in classifying topoi, in *Category theory 1991*, C.M.S. Conf. Proc. vol. 13 (Amer. Math. Soc., 1992), 1–31; MR 94h:03069.

[41] A. Bagchi, The Lee identities in topoi, I, *J. Pure Appl. Alg.* 120 (1997), 143–159; MR 98k:18003.

[42] E. S. Bainbridge, P. J. Freyd, A. Scedrov and P. J. Scott, Functorial polymorphism, *Theoret. Comput. Sci.* 70 (1990), 35–64; errata *ibid.* 71 (1990), 431; MR 91g:03129; also in *Logical foundations of functional programming* (Addison–Wesley, 1990), 315–327.

[43] D. Ballard and W. Boshuck, Definability and descent, *J. Symbolic Logic* 63 (1998), 372–378; MR 99f:03092.

[44] B. Banaschewski, Recovering a space from its Banach sheaves, in *Categorical topology*, Lecture Notes in Math. vol. 719 (1979), 1–12; MR 80m:18004.

[45] B. Banaschewski, The duality of distributive continuous lattices, *Canad. J. Math.* 32 (1980), 385–394; MR 81m:54072.
[46] B. Banaschewski, Extension of invariant linear functionals: Hahn–Banach in the topos of M-sets, *J. Pure Appl. Alg.* 17 (1980), 227–248; MR 82c:46801.
[47] B. Banaschewski, Injective sheaves of abelian groups: a counterexample, *Canad. J. Math.* 32 (1980), 1518–1521; MR 82f:18015a.
[48] B. Banaschewski, When are divisible abelian groups injective?, *Quaest. Math.* 4 (1981), 285–307; MR 83e:20063.
[49] B. Banaschewski, Coherent frames, in *Continuous lattices*, Lecture Notes in Math. vol. 871 (1981), 1–11.
[50] B. Banaschewski, Injectivity and modelling in the Blass topos, *J. Pure Appl. Alg.* 49 (1987), 1–10; MR 89a:18006.
[51] B. Banaschewski, Another look at the localic Tychonoff theorem, *Comment. Math. Univ. Carolinae* 29 (1988), 647–656; MR 90a:06002.
[52] B. Banaschewski, Radical ideals and coherent frames, *Comment. Math. Univ. Carolinae* 37 (1996), 349–370; MR 97g:13002.
[53] B. Banaschewski, A uniform view of localic realcompactness, *J. Pure Appl. Alg.* 143 (1999), 49–68; MR 01i:54023.
[54] B. Banaschewski and K. R. Bhutani, Boolean algebras in a localic topos, *Math. Proc. Camb. Philos. Soc.* 100 (1986), 43–55; MR 87f:18003.
[55] B. Banaschewski and G. C. L. Brümmer, Thoughts on the Cantor–Bernstein theorem, *Quaest. Math.* 9 (1986), 1–27; MR 87i:04001.
[56] B. Banaschewski and R. Harting, Lattice aspects of radical ideals and choice principles, *Proc. Lond. Math. Soc.* (3) 50 (1985), 385–404; MR 86g:16047.
[57] B. Banaschewski and G. H. Moore, The dual Cantor–Bernstein theorem and the partition principle, *Notre Dame J. Formal Logic* 31 (1990), 375–381; MR 91k:03128.
[58] B. Banaschewski and C. J. Mulvey, Stone–Čech compactification of locales, I, *Houston J. Math.* 6 (1980), 301–312; MR 82b:06010.
[59] B. Banaschewski and C. J. Mulvey, Stone–Čech compactification of locales, II, *J. Pure Appl. Alg.* 33 (1984), 107–122; MR 86d:06009.
[60] B. Banaschewski and C. J. Mulvey, A constructive proof of the Stone–Weierstrass theorem, *J. Pure Appl. Alg.* 116 (1997), 25–40; MR 98d:46028.
[61] B. Banaschewski and C. J. Mulvey, The spectral theory of commutative $C*$-algebras: the constructive spectrum, *Quaest. Math.* 23 (2000), 425–464; MR 02b:46112a.
[62] B. Banaschewski and C. J. Mulvey, The spectral theory of commutative $C*$-algebras: the constructive Gelfand–Mazur theorem, *Quaest. Math.* 23 (2000), 465–488; MR 02b:46112b.
[63] B. Banaschewski and S. B. Niefield, Projective and supercoherent frames, *J. Pure Appl. Alg.* 70 (1991), 45–51; MR 92b:06034.
[64] B. Banaschewski and A. Pultr, Variants of openness, *Appl. Categ. Struct.* 2 (1994), 331–350; MR 95j:06012.
[65] B. Banaschewski and A. Pultr, Booleanization, *Cahiers Top. Géom. Diff. Catégoriques* 37 (1996), 41–60; MR 97b:06006.
[66] B. Banaschewski and J. J. C. Vermeulen, Polynomials and radical ideals, *J. Pure Appl. Alg.* 113 (1996), 219–227; MR 98c:13004.
[67] H. Barendregt, *The lambda calculus: its syntax and semantics*, Studies in Logic and the Foundations of Math. vol. 103 (North-Holland, 1984); MR 86a:03012.
[68] M. Barr, Exact categories, in *Exact categories and categories of sheaves*, Lecture Notes in Math. vol. 236 (Springer-Verlag, 1971), 1–120.
[69] M. Barr, The point of the empty set, *Cahiers Top. Géom. Diff.* 13 (1972), 357–368; MR 48/2216.

[70] M. Barr, Right exact functors, *J. Pure Appl. Alg.* 4 (1974), 1–8; MR 51/10439.

[71] M. Barr, Toposes without points, *J. Pure Appl. Alg.* 5 (1974), 265–280; MR 53/13354.

[72] M. Barr, *$*$-Autonomous categories*, Lecture Notes in Math. vol. 752 (Springer–Verlag, 1979); MR 81h:18008.

[73] M. Barr, Abstract Galois theory, *J. Pure Appl. Alg.* 19 (1980), 21–42; MR 81m:18001.

[74] M. Barr, Abstract Galois theory, II, *J. Pure Appl. Alg.* 25 (1982), 227–247; MR 83j:18001.

[75] M. Barr, Representation of categories, *J. Pure Appl. Alg.* 41 (1986), 113–137; MR 87h:18005.

[76] M. Barr, Fuzzy set theory and topos theory, *Canad. Math. Bull.* 29 (1986), 501–508; MR 88a:03129.

[77] M. Barr, Models of sketches, *Cahiers Top. Géom. Diff. Catégoriques* 27 (1986), 93–107; MR 87h:18009.

[78] M. Barr, On categories with effective unions, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 19–35; MR 89k:18010.

[79] M. Barr, Models of Horn theories, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 1–7; MR 90k:18006.

[80] M. Barr, Accessible categories and models of linear logic, *J. Pure Appl. Alg.* 69 (1990), 219–232; MR 92a:03043.

[81] M. Barr and R. Diaconescu, Atomic toposes, *J. Pure Appl. Alg.* 17 (1980), 1–24; MR 81e:18013.

[82] M. Barr and R. Diaconescu, On locally simply connected toposes and their fundamental groups, *Cahiers Top. Géom. Diff.* 22 (1981), 301–314; MR 83e:18002.

[83] M. Barr and M. Makkai, On representations of Grothendieck toposes, *Canad. J. Math.* 39 (1987), 168–221; MR 88f:18005.

[84] M. Barr and R. Paré, Molecular toposes, *J. Pure Appl. Alg.* 17 (1980), 127–152; MR 81d:18010.

[85] M. Barr and M. C. Pedicchio, $\mathbf{Top}^{\mathrm{op}}$ is a quasi-variety, *Cahiers Top. Géom. Diff. Catégoriques* 36 (1995), 3–10; MR 96c:08013.

[86] M. Barr and M. C. Pedicchio, Topological spaces and quasi-varieties, *Appl. Categ. Struct.* 4 (1996), 81–85; MR 97d:54018.

[87] M. Barr and C. Wells, *Toposes, triples and theories*, Grundl. Math. Wissenschaften no. 278 (Springer–Verlag, 1985); MR 86f:18001.

[88] M. Barr and C. Wells, *Category theory for computing science* (Prentice-Hall, 1990); MR 92g:18001.

[89] M. Barr and C. Wells, On the limitations of sketches, *Canad. Math. Bull.* 35 (1992), 287–294; MR 93i:18009.

[90] M. J. Beeson, *Foundations of constructive mathematics*, Ergebnisse der Math., 3. Folge, Band 6 (Springer–Verlag, 1985); MR 86k:03055.

[91] J. L. Bell, Isomorphism of structures in $\mathcal{S}$-toposes, *J. Symbolic Logic* 46 (1981), 449–459; MR 83a:03068.

[92] J. L. Bell, Categories, toposes and sets, *Synthèse* 51 (1982), 293–337; MR 83j:03109.

[93] J. L. Bell, Some aspects of the category of subobjects of constant objects in a topos, *J. Pure Appl. Alg.* 24 (1982), 245–259; MR 83m:18004.

[94] J. L. Bell, On the strength of the Sikorski extension theorem for Boolean algebras, *J. Symbolic Logic* 48 (1983), 841–846; MR 86b:03061.

[95] J. L. Bell, *Toposes and local set theories: an introduction*, Oxford Logic Guides no. 14 (Oxford Univ. Press, 1988); MR 90e:18002.

[96] J. L. Bell, Zorn's lemma and complete Boolean algebras in intuitionistic type theories, *J. Symbolic Logic* 62 (1997), 1265–1279; MR 99g:03062.

[97] J. L. Bell, *A primer of infinitesimal analysis* (Cambridge Univ. Press, 1998); MR 99h:00002.

[98] J. L. Bell and S. Gebellato, Precovers, modalities and universal closure operations in a topos, *Math. Logic Quart.* 42 (1996), 289–299; MR 98c:03136.

[99] J. Bénabou, Introduction to bicategories, in *Reports of the Midwest category seminar*, Lecture Notes in Math. vol. 47 (Springer–Verlag, 1967), 1–77; MR 36/3841.

[100] J. Bénabou, Théories relatives à un corpus, *C. R. Acad. Sci. Paris* 281 (1975), A831–834; MR 52/13990.

[101] J. Bénabou, Fibrations petites et localement petites, *C. R. Acad. Sci. Paris* 281 (1975), A897–900; MR 52/13991.

[102] J. Bénabou, Fibred categories and the foundations of naive category theory, *J. Symbolic Logic* 50 (1985), 10–37; MR 87h:18001.

[103] J. Bénabou, Some remarks on 2-categorical algebra, I, *Bull. Soc. Math. Belg.* (A) 41 (1989), 127–194; MR 91a:18008.

[104] J. Bénabou, Some remarks on free monoids in a topos, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 20–29; MR 93h:18008.

[105] J. Bénabou, Some geometric aspects of the calculus of fractions, *Appl. Categ. Struct.* 4 (1996), 139–165; MR 97g:18007.

[106] J. Bénabou and B. Loiseau, Orbits and monoids in a topos, *J. Pure Appl. Alg.* 92 (1994), 29–54; MR 94k:18002.

[107] R. Betti, Categorie variabili, *Rend. Sem. Mat. Fis. Milano* 57 (1987), 483–518; MR 91b:18010.

[108] R. Betti, Change of base for locally internal categories, *J. Pure Appl. Alg.* 61 (1989), 233–242; MR 91a:18007.

[109] R. Betti, An elementary characterizaton of presheaves categories, *Rend. Circ. Mat. Palermo* Suppl. (2) 24 (1990), 257–266; MR 92c:18014.

[110] R. Betti, On internal topologies and the inverse image, *Rend. Circ. Mat. Palermo* Suppl. (2) 99 (1992), 339–352; MR 93k:18009.

[111] R. Betti, Sheaves in cocomplete categories, in *Category theory 1991*, C.M.S. Conf. Proc. vol. 13 (Amer. Math. Soc., 1992), 43–54; MR 93j:18006.

[112] R. Betti, Formal theory of internal categories, *Le Matematiche* 51 (1996), Suppl. 35–52; MR 98h:18001.

[113] R. Betti, Weak equivalence of internal categories, *Appl. Categ. Struct.* 8 (2000), 307–316; MR 01c:18010.

[114] R. Betti, Descent theory of locally internal categories, *J. Pure Appl. Alg.* 154 (2000), 15–26; MR 01f:18021.

[115] R. Betti and A. Carboni, Cauchy completion and the associated sheaf, *Cahiers Top. Géom. Diff.* 23 (1982), 243–256; MR 84g:18026.

[116] K. R. Bhutani, Injectivity and injective hulls of abelian groups in a localic topos, *Bull. Austral. Math. Soc.* 37 (1988), 43–59; MR 89a:18018.

[117] K. R. Bhutani, Abelian groups in a topos of sheaves: torsion and essential extensions, *Internat. J. Math. Math. Sci.* 12 (1989), 89–98; MR 90a:18008.

[118] K. R. Bhutani, Stability of abelian groups in a topos of sheaves, *J. Pure Appl. Alg.* 68 (1990), 47–54; MR 91m:18007.

[119] L. Biondo and G. C. Meloni, Locali 'étalé' nel contesto dei topoi, *Boll. Un. Mat. Ital.* (7) 8-A (1994), 157–167; MR 95c:18004.

[120] G. J. Bird, G. M. Kelly, A. J. Power and R. H. Street, Flexible limits for 2-categories, *J. Pure Appl. Alg.* 61 (1989), 1–27; MR 91a:18009.

[121] A. R. Blass, Exact functors and measurable cardinals, *Pacific J. Math.* 63 (1976), 335–346; MR 54/2466.

[122] A. R. Blass, Two closed categories of filters, *Fund. Math.* 94 (1977), 129–143; MR 56/3083.

[123] A. R. Blass, Injectivity, projectivity and the axiom of choice, *Trans. Amer. Math. Soc.* 255 (1979), 31–59; MR 81j:04004.

[124] A. R. Blass, Words, free algebras and coequalizers, *Fund. Math.* 117 (1983), 117–160; MR 85d:03127.

[125] A. R. Blass, Cohomology detects failures of the axiom of choice, *Trans. Amer. Math. Soc.* 279 (1983), 257–269; MR 84h:03116.

[126] A. R. Blass, The interaction between category theory and set theory, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 5–29; MR 85m:03045.

[127] A. R. Blass, Prime ideals yield almost maximal ideals, *Fund. Math.* 127 (1986), 57–66; MR 88d:03099.

[128] A. R. Blass, Functions on universal algebras, *J. Pure Appl. Alg.* 42 (1986), 25–28; MR 87i:03134.

[129] A. R. Blass, Well-ordering and induction in intuitionistic logic and topoi, in *Mathematical logic and theoretical computer science*, Lecture Notes Pure Appl. Math. vol. 106 (Marcel Dekker, 1987), 29–48; MR 89c:03095.

[130] A. R. Blass, Classifying topoi and the axiom of infinity, *Algebra Universalis* 26 (1989), 341–345; MR 91a:03117.

[131] A. R. Blass, Geometric invariance of existential fixed-point logic, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 9–22; MR 90g:03066.

[132] A. R. Blass, An induction principle and pigeonhole principles for K-finite sets, *J. Symbolic Logic* 60 (1995), 1186–1193; MR 96k:03140.

[133] A. R. Blass and A. Ščedrov, Boolean classifying topoi, *J. Pure Appl. Alg.* 28 (1983), 15–30; MR 84e:03078.

[134] A. R. Blass and A. Ščedrov, Classifying topoi and finite forcing, *J. Pure Appl. Alg.* 28 (1983), 111–140; MR 84g:03035.

[135] A. R. Blass and A. Ščedrov, Small decidable sheaves, *J. Symbolic Logic* 51 (1986), 726–731; MR 87i:03124.

[136] A. R. Blass and A. Scedrov, *Freyd's models for the independence of the axiom of choice*, Memoirs Amer. Math. Soc. no. 404 (1989); MR 90d:03097.

[137] A. R. Blass and A. Scedrov, Complete topoi representing models of set theory, *Ann. Pure Appl. Logic* 57 (1992), 1–26; MR 93d:03056.

[138] A Boileau and A. Joyal, La logique des topos, *J. Symbolic Logic* 46 (1981), 6–16; MR 82a:03063.

[139] F. Borceux, When is Ω a cogenerator in a topos?, *Cahiers Top. Géom. Diff.* 16 (1975), 3–15; MR 52/3277.

[140] F. Borceux, Topologies sur les topos de faisceaux, *Ann. Soc. Sci. Bruxelles* 94 (1980), 163–165; MR 82e:54013.

[141] F. Borceux, Algebraic localizations and elementary toposes, *Cahiers Top. Géom. Diff.* 21 (1980), 393–401; MR 82g·18002.

[142] F. Borceux, Sheaves of algebras for a commutative theory, *Ann. Soc. Sci. Bruxelles* 95 (1981), 3–19; MR 83c:18006.

[143] F. Borceux, Higher order sheaves and Banach modules, in *categorical topology*, Lecture Notes in Math. vol. 915 (Springer–Verlag, 1982), 40–49; MR 83f:46079.

[144] F. Borceux, On algebraic localizations, in *Category theory*, Lecture Notes in Math. vol. 962 (Springer–Verlag, 1982), 13–22; MR 84c:18007.

[145] F. Borceux, A propos d'un théorème de Barr, *Ann. Soc. Sci. Bruxelles* 97 (1983), 41–58; MR 86c:18002.

[146] F. Borceux, Examples of quantales in topos theory, *Ann. Soc. Sci. Bruxelles* 101 (1987), 61–67; MR 90a:18003.

[147] F. Borceux, *Handbook of categorical algebra* (3 volumes) (Cambridge Univ. Press, 1994); MR 96g:18001.

[148] F. Borceux and R. Cruciani, Sheaves on a quantale, *Cahiers Top. Géom. Diff. Catégoriques* 34 (1993), 209–228; MR 94h:06017.

[149] F. Borceux and R. Cruciani, Skew Ω-sets coincide with Ω-posets, *Cahiers Top. Géom. Diff. Catégoriques* 39 (1998), 205–220; MR 00b:18003.

[150] F. Borceux and B. J. Day, Universal algebra in a closed category, *J. Pure Appl. Alg.* 16 (1980), 133–147; MR 81f:18023.

[151] F. Borceux and G. M. Kelly, On locales of localizations, *J. Pure Appl. Alg.* 46 (1987), 1–34; MR 88k:18004.

[152] F. Borceux and M. Korostenski, Open localizations, *J. Pure Appl. Alg.* 74 (1991), 229–238; MR 92k:18003.

[153] F. Borceux and M. C. Pedicchio, De Morgan's and strong De Morgan's laws in a topos of sheaves, *Acta Univ. Carolin. Math. Phys.* 28 (1987), 9–12; MR 89e:18007.

[154] F. Borceux and M. C. Pedicchio, On primeness and maximality of filters, *Cahiers Top. Géom. Diff. Catégoriques* 30 (1989), 167–181; MR 90g:06023.

[155] F. Borceux and M. C. Pedicchio, Prime elements in a localic topos, *Algebra Universalis* 26 (1989), 166–177; MR 90g:06021.

[156] F. Borceux and M. C. Pedicchio, A characterization of quasi-toposes, *J. Algebra* 139 (1991), 505–526; MR 92e:18002.

[157] F. Borceux and M. C. Pedicchio, Left exact presheaves on a small pretopos, *J. Pure Appl. Alg.* 135 (1999), 9–22; MR 99j:18012.

[158] F. Borceux, M. C. Pedicchio and F. Rossi, Boolean algebras in a localic topos, *J. Pure Appl. Alg.* 68 (1980), 55–65; MR 92b:18002.

[159] F. Borceux, C. Quinteiro and J. Rosický, A theory of enriched sketches, *Theory Appl. Categ.* 4 (1998), 47–72; MR 99i:18009.

[160] F. Borceux, H. Simmons and G. Van Den Bossche, A sheaf representation for modules with applications to Gelfand rings, *Proc. London Math. Soc.* (3) 48 (1984), 230–246; MR 85j:16061.

[161] F. Borceux and G. Van Den Bossche, Recovering a frame from its sheaves of algebras, *J. Pure Appl. Alg.* 28 (1983), 141–154; MR 85e:18009.

[162] F. Borceux and G. Van Den Bossche, *Algebra in a localic topos with applications to ring theory*, Lecture Notes in Math. vol. 1038 (Springer–Verlag, 1983); MR 85h:18005.

[163] F. Borceux and G. Van Den Bossche, Structure des topologies d'un topos, *Cahiers Top. Géom. Diff.* 25 (1984), 37–39; MR 86a:18006.

[164] F. Borceux and G. Van Den Bossche, The structure of algebraic localizations, *Rev. Roum. Math. Pures Appl.* 31 (1986), 367–370; MR 88c:18001.

[165] F. Borceux and G. Van Den Bossche, Quantales and their sheaves, *Order* 3 (1986), 61–87; MR 87k:18012.

[166] F. Borceux and B. Veit, Continuous Grothendieck topologies, *Ann. Soc. Sci. Bruxelles* 100 (1986), 31–42; MR 88e:18006.

[167] F. Borceux and B. Veit, On the left exactness of orthogonal reflections, *J. Pure Appl. Alg.* 49 (1987), 33–42; MR 88m:18005.

[168] F. Borceux and B. Veit, Subobject classifier for algebraic structures, *J. Algebra* 112 (1988), 306–314; MR 89g:18006.

[169] R. Börger, Disjointness and related properties of coproducts, *Acta Univ. Carolin. Math. Phys.* 35 (1994), 43–63; MR 96k:18001.

[170] D. Bourn, Ditopos, *C. R. Acad. Sci. Paris* 279 (1974), A731–732 and A911–913; MR 51/639.

[171] D. Bourn, Une autre propriété universelle pour le champ associé, *Cahiers Top. Géom. Diff.* 21 (1980), 403–410; MR 82e:18005.

[172] S. Bozapalides, Some remarks on lax-presheafs, *Ill. J. Math.* 24 (1980), 676–680; MR 81m:18006.
[173] S. Bozzi and G. C. Meloni, Completezza propositionale per l'operatore modale 'accade localmente che', *Boll. Un. Mat. Ital.* A (5) 14 (1977), 489–497; MR 57/9491.
[174] S. Bozzi and G. C. Meloni, Representation of Heyting algebras with covering and propositional intuitionistic logic with local operator, *Boll. Un. Mat. Ital.* A (5) 17 (1980), 436–442; MR 82k:03099.
[175] L. Breen, Bitorseurs et cohomologie non abélienne, in *The Grothendieck Festschrift*, vol. I, Progress in Math. vol. 86 (Birkhäuser–Verlag, 1990), 401–476; MR 92m:18019.
[176] O. P. Bruno, Logical opens of exponential objects, *Cahiers Top. Géom. Diff. Catégoriques* 26 (1985), 311–323; MR 87f:58002.
[177] O. P. Bruno, Vector fields on R^R in well adapted models of synthetic differential geometry, *J. Pure Appl. Alg.* 45 (1987), 1–14; MR 88m:18007.
[178] A. Bucalo and G. Rosolini, Repleteness and the associated sheaf, *J. Pure Appl. Alg.* 127 (1998), 147–151; MR 99k:18005.
[179] I. Bucur, Quelques propriétés des topos, *Rev. Roum. Math. Pures Appl.* 23 (1978), 161–167; MR 58/5855.
[180] M. C. Bunge, Relative functor categories and categories of algebras, *J. Algebra* 11 (1969), 64–101; MR 38/4536.
[181] M. C. Bunge, Topos theory and Souslin's hypothesis, *J. Pure Appl. Alg.* 4 (1974), 159–187; MR 51/2908; corrigenda *ibid.* 14 (1979), 115–116; MR 80d:18006.
[182] M. C. Bunge, Internal presheaves toposes, *Cahiers Top. Géom. Diff.* 18 (1977), 291–330; MR 57/411.
[183] M. C. Bunge, Stack completion and Morita equivalence for categories in a topos, *Cahiers Top. Géom. Diff.* 20 (1979), 401–436; MR 81f:18018.
[184] M. C. Bunge, Sheaves and prime model extensions, *J. Algebra* 68 (1981), 79–96; MR 83d:03043.
[185] M. C. Bunge, Models of differential algebra in the context of synthetic differential geometry, *Cahiers Top. Géom. Diff.* 22 (1981), 31–44; MR 83a:12033.
[186] M. C. Bunge, On the transfer of an abstract Nullstellensatz, *Commun. Alg.* 10 (1982), 1891–1906; MR 84c:03074.
[187] M. C. Bunge, Synthetic aspects of C^∞-mappings, *J. Pure Appl. Alg.* 28 (1983), 41–63; MR 84h:58019.
[188] M. C. Bunge, Toposes in logic and logic in toposes, *Topoi* 3 (1984), 13–22; MR 86e:03060.
[189] M. C. Bunge, An application of descent to a classification theorem for toposes, *Math. Proc. Camb. Philos. Soc.* 107 (1990), 59–79; MR 90k:18002.
[190] M. C. Bunge, Classifying toposes and fundamental localic groupoids, in *Category theory 1991*, C.M.S. Conf. Proc. vol. 13 (Amer. Math. Soc., 1992), 75–96; MR 93j:18001.
[191] M. C. Bunge, Universal covering localic toposes, *C. R. Math. Rep. Acad. Sci. Canada* 24 (1992), 245–250; MR 93m:18007.
[192] M. C. Bunge, Cosheaves and distributions on toposes, *Alg. Universalis* 34 (1995), 469–484; MR 98g:18001.
[193] M. C. Bunge and A. Carboni, The symmetric topos, *J. Pure Appl. Alg.* 105 (1995), 233–249; MR 96i:18004.
[194] M. C. Bunge and E. J. Dubuc, Archimedean local C^∞-rings and models of synthetic differential geometry, *Cahiers Top. Géom. Diff. Catégoriques* 27 (1986), 3–22; MR 88c:58004.

[195] M. C. Bunge and E. J. Dubuc, Local concepts in SDG and germ representability, in *Mathematical logic and theoretical computer science*, Lecture Notes Pure Appl. Math. vol. 106 (Marcel Dekker, 1987), 39–158; MR 89f:58005.

[196] M. C. Bunge and J. Funk, Constructive theory of the lower power locale, *Math. Struct. Comput. Sci.* 6 (1996), 1–15; MR 98f:18004.

[197] M. C. Bunge and J. Funk, Spreads and the symmetric topos, *J. Pure Appl. Alg.* 113 (1996), 1–38; MR 98f:18003.

[198] M. C. Bunge and J. Funk, Spreads and the symmetric topos, II, *J. Pure Appl. Alg.* 130 (1998), 49–84; MR 99k:18006.

[199] M. C. Bunge and J. Funk, On a bicomma object condition for KZ-doctrines, *J. Pure Appl. Alg.* 143 (1999), 69–105; MR 01b:18006.

[200] M. C. Bunge, J. Funk, M. Jibladze and T. Streicher, Distribution algebras and duality, *Adv. Math.* 156 (2000), 133–155; MR 02a:18005.

[201] M. C. Bunge and F. Gago Couso, Synthetic aspects of C^∞-mappings, II: Mather's theorem for infinitesimally represented germs, *J. Pure Appl. Alg.* 55 (1988), 213–250; MR 90b:58022.

[202] M. C. Bunge and M. Heggie, Synthetic calculus of variations, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 30–62; MR 85m:18004.

[203] M. C. Bunge and I. Moerdijk, On the construction of the Grothendieck fundamental group of a topos by paths, *J. Pure Appl. Alg.* 116 (1997), 99–113; MR 98b:18007.

[204] M. C. Bunge and S. B. Niefield, Exponentiability and single universes, *J. Pure Appl. Alg.* 148 (2000), 217–250; MR 01f:18020.

[205] M. C. Bunge and R. Paré, Stacks and equivalence of indexed categories, *Cahiers Top. Géom. Diff.* 20 (1979), 373–399; MR 81f:18017.

[206] M. C. Bunge and G. E. Reyes, Boolean spectra and model completions, *Fund. Math.* 113 (1981), 165–173; MR 83a:03069.

[207] M. C. Bunge and P. Sawyer, On connections, geodesics and sprays in synthetic differential geometry, *Cahiers Top. Géom. Diff.* 25 (1984), 221–258; MR 86k:58011.

[208] C. W. Burden, The Hahn–Banach theorem in a category of sheaves, *J. Pure Appl. Alg.* 17 (1980), 25–34; MR 81h:46096.

[209] C. W. Burden and C. J. Mulvey, Banach spaces in categories of sheaves, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 169–196; MR 81h:46095.

[210] C. Butz, Syntax and semantics of the logic $\mathcal{L}^{\lambda}_{\omega\omega}$, *Notre Dame J. Formal Logic* 38 (1997), 374–384; MR 00a:03111.

[211] C. Butz, A topological completeness theorem, *Arch. Math. Logic* 38 (1999), 79–101; MR 00i:03092.

[212] C. Butz and P. T. Johnstone, Classifying toposes for first-order theories, *Ann. Pure Appl. Logic* 91 (1998), 33–58; MR 99b:03081.

[213] C. Butz and I. Moerdijk, Représentation de topos par des espaces topologiques, *C. R. Acad. Sci. Paris* Sér. I 325 (1997), 461–464; MR 00b:18004.

[214] C. Butz and I. Moerdijk, Representing topoi by topological groupoids, *J. Pure Appl. Alg.* 130 (1998), 223–235; MR 99i:18005.

[215] C. Butz and I. Moerdijk, Topological representation of sheaf cohomology of sites, *Compositio Math.* 118 (1999), 217–233; MR 00g:14027.

[216] C. Butz and I. Moerdijk, An elementary definability theorem for first order logic, *J. Symbolic Logic* 64 (1999), 1028–1036; MR 01h:03055.

[217] X. Caicedo, Logic of sheaves of structures (Spanish), *Rev. Acad. Colombiana Cienc. Exact. Fis. Natur.* 19 (1995), 569–586; MR 98g:03149.

[218] A. Carboni, Categorie di relazioni, *Ist. Lombardo Accad. Sci. Lett. Rend.* A 110 (1976), 342–350; MR 58/22228.

[219] A. Carboni, Analisi non-standard e topos, *Rend. Ist. Mat. Univ. Trieste* 14 (1982), 1–16; MR 85k:18005.

[220] A. Carboni, Some free constructions in realizability and proof theory, *J. Pure Appl. Alg.* 103 (1995), 117–148; MR 96j:03088.

[221] A. Carboni and R. Celia Magno, The free exact category on a left exact one, *J. Austral. Math. Soc.* A 33 (1982), 295–301; MR 84g:18017.

[222] A. Carboni, P. J. Freyd and A. Scedrov, A categorical approach to realizability and polymorphic types, in *Mathematical foundations of programming language semantics*, Lecture Notes in Computer Science vol. 298 (Springer–Verlag, 1988), 23–42; MR 89g:68403.

[223] A. Carboni and M. Galuzzi, Anelli regolari e topoi, *Atti Accad. Sci. Torino* 111 (1977), 227–231; MR 57/6143.

[224] A. Carboni and P. T. Johnstone, Connected limits, familial representability and Artin glueing, *Math. Struct. Comput. Sci.* 5 (1995), 441–459; MR 98f:18005.

[225] A. Carboni, G. M. Kelly, D. R. B. Verity and R. J. Wood, A 2-categorical approach to change of base and geometric morphisms, II, *Theory Appl. Categ.* 4 (1998), 82–136; MR 99e:18007.

[226] A. Carboni, G. M. Kelly and R. J. Wood, A 2-categorical approach to change of base and geometric morphisms, I, *Cahiers Top. Géom. Diff. Catégoriques* 32 (1991), 47–95; MR 92j:18003.

[227] A. Carboni, S. G. Lack and R. F. C. Walters, Introduction to extensive and distributive categories, *J. Pure Appl. Alg.* 84 (1993), 145–158; MR 93k:18004.

[228] A. Carboni and S. Mantovani, An elementary characterization of categories of separated objects, *J. Pure Appl. Alg.* 89 (1993), 63–92; MR 94k:18013.

[229] A. Carboni and G. C. Meloni, Costruzione algebrica dello spazio étalé, *Boll. Un. Mat. Ital.* (4) 12 (1975), 192–197; MR 52/13998.

[230] A. Carboni and G. Rosolini, Locally cartesian closed exact completions, *J. Pure Appl. Alg.* 154 (2000), 103–116; MR 01f:18009.

[231] A. Carboni and E. M. Vitale, Regular and exact completions, *J. Pure Appl. Alg.* 125 (1998), 79–116; MR 99b:18003.

[232] A. Carboni and R. F. C. Walters, Cartesian bicategories I, *J. Pure Appl. Alg.* 49 (1987), 11–32; MR 88k:18009.

[233] S. Carel, $I_{n,m}$-objekt in einem Elementartopos, *Bul. Univ. Braşov* C 30 (1988), 65–67; MR 89i:18002.

[234] J. C. Carrega, The categories **Set** H and **Fuz** H, *Fuzzy Sets Syst.* 3 (1983), 327–332; MR 84i:18002.

[235] H. Cartan, Théorie des filtres, *C. R. Acad. Sci. Paris* 205 (1937), 595–598.

[236] P. Cartier, Catégories, logique et faisceaux: modèles de la théorie des ensembles, *Sém. Bourbaki*, exposé 513 (1978); MR 80m:03111.

[237] J. Cartmell, Generalised algebraic theories and contextual categories, *Ann. Pure Appl. Logic* 32 (1986), 209–243; MR 88j:03046.

[238] J. Chapman and F. Rowbottom, *Relative category theory and geometric morphisms: a logical approach*, Oxford Logic Guides no. 16 (Oxford Univ. Press, 1991); MR 93i:18004.

[239] B. P. Chisala and M. M. Mawanda, Counting measure for Kuratowski-finite parts and decidability, *Cahiers Top. Géom. Diff. Catégoriques* 32 (1991), 345–353; MR 93f:03044.

[240] G. Choquet, Convergences, *Ann. Univ. Grenoble, Sect. Sci. Math. Phys. (N.S.)* 23 (1948), 57–112; MR 10–53.

[241] A. Church, A set of postulates for the foundation of logic, *Ann. Math.* (2) 33 (1932), 346–366.

[242] J. R. B. Cockett, On the decidability of objects in a locos, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 23–46; MR 90f:03101.
[243] J. R. B. Cockett, List-arithmetic distributive categories: locoi, *J. Pure Appl. Alg.* 66 (1990), 1–29; MR 91k:03152.
[244] J. R. B. Cockett, Introduction to distributive categories, *Math. Struct. Comput. Sci.* 3 (1993), 277–307; MR 94m:18002.
[245] J. C. Cole, Categories of sets and models of set theory, in *Proc. Bertrand Russell Memorial Logic Conf., Uldum 1971* (Leeds, 1973), 351–399; MR 50/9584.
[246] J. C. Cole, Classifying topoi, in *Algebraische Modelle, Kategorien und Gruppoide*, Studien zur Algebra und ihre Anwendungen Band 7 (Akademie–Verlag, 1979), 5–20; MR 81e:18008.
[247] F. Conduché, Au sujet de l'existence d'adjoints à droite aux foncteurs 'image réciproque' dans la catégorie des catégories, *C. R. Acad. Sci. Paris* 275 (1972), A891–894; MR 46/9136.
[248] J. H. Conway, *On numbers and games*, L.M.S. Monographs no. 6 (Academic Press, 1976); MR 56/8365.
[249] T. Coquand, An intuitionistic proof of Tychonoff's theorem, *J. Symbolic Logic* 57 (1992), 28–32; MR 93c:03076.
[250] T. Coquand, S. Sadocco, G. Sambin and J. M. Smith, Formal topologies on the set of first-order formulae, *J. Symbolic Logic* 65 (2000), 1183–1192; MR 01j:03112.
[251] W. H. Cornish, The Chinese remainder theorem and sheaf representations, *Fund. Math.* 96 (1977), 177–187; MR 58/444.
[252] M. F. Coste, Construction d'un modèle booléen de la théorie des ensembles à partir d'un topos booléen, *C. R. Acad. Sci. Paris* 278 (1974), A1073–1076; MR 49/5114.
[253] M. F. Coste and M. Coste, The generic model of an ϵ-stable geometric extension of rings is of line type, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 29–36; MR 80m:03112.
[254] M. F. Coste and M. Coste, Topologies for real algebraic geometry, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 37–100; MR 80m:14002.
[255] M. F. Coste and M. Coste, Le spectre étale réel d'un anneau est spatial, *C. R. Acad. Sci. Paris* 290 (1980), A91–94; MR 81c:14001.
[256] M. F. Coste-Roy and M. Coste, Le topos étale réel d'un anneau, *Cahiers Top. Géom. Diff.* 22 (1981), 19–24; MR 82h:14016.
[257] M. F. Coste, M. Coste and L. Mahé, Contributions to the study of the natural number object in elementary topoi, *J. Pure Appl. Alg.* 17 (1980), 35–68; MR 81h:18005.
[258] M. Coste, Localization, spectra and sheaf representation, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 212–238; MR 83c:18002.
[259] M. Coste and G. Michon, Petits et gros topos en géométrie algébrique, *Cahiers Top. Géom. Diff.* 22 (1981), 25–30; MR 82g:14024.
[260] J. Coulon and J. L. Coulon, Is the category JTF a topos?, *Fuzzy Sets Syst.* 27 (1988), 11–19; MR 89h:18003.
[261] J. Coulon and J. L. Coulon, Morphisms in the category JTF^∞: weak classifier of monomorphisms, *Portugal. Math.* 45 (1988), 29–47; MR 89h:18004.
[262] J. Coulon and J. L. Coulon, Injectivity in the quasitopos of separated presheaves on a complete Heyting algebra, *Portugal. Math.* 47 (1990), 227–241; MR 92e:03085b.

[263] J. Coulon and J. L. Coulon, About some categories of Ω-valued sets, *J. Math. Anal. Appl.* 154 (1991), 273–278; MR 92e:03085a.
[264] M. Crabbé, Prelogic of logoi, *Studia Logica* 35 (1976), 219–226; MR 55/10248.
[265] S. E. Crans, Quillen closed model structures for sheaves, *J. Pure Appl. Alg.* 101 (1995), 35–57; MR 97f:18005.
[266] R. L. Crole, *Categories for types* (Cambridge Univ. Press, 1993); MR 95h:18007.
[267] D. Ćubrić, On the semantics of the universal quantifier, *Ann. Pure Appl. Logic* 87 (1997), 209–239; MR 99a:03009.
[268] P. L. Curien and A. Obtułowicz, Partiality, cartesian closedness and toposes, *Inform. Comput.* 80 (1989), 50–95; MR 90b:18010b.
[269] H. B. Curry, Grundlagen der kombinatorischen Logik, *Amer. J. Math.* 52 (1930), 509–536.
[270] A. Daigneault, Lawvere's elementary theories and polyadic and cylindric algebras, *Fund. Math.* 66 (1970), 307–328; MR 41/5205.
[271] D. van Dalen and F. J. de Vries, Intuitionistic free abelian groups, *Z. Math. Logik Grundl. Math.* 34 (1988), 3–12; MR 89c:03096.
[272] C. Dapueto, Interpretations in an elementary topos, *Rend. Mat.* (7) 2 (1982), 531–545; MR 84f:03056.
[273] B. A. Davey, Sheaf spaces and sheaves of universal algebras, *Math. Z.* 134 (1973), 275–290; MR 48/8345.
[274] R. C. Davis, Toposes, monoid actions and universal coalgebra, *Proc. Amer. Math. Soc.* 98 (1986), 547–552; MR 87m:18005.
[275] B. J. Day, An adjoint-functor theorem over topoi, *Bull. Austral. Math. Soc.* 15 (1976), 381–394; MR 61/450.
[276] B. J. Day and G. M. Kelly, On topological quotient maps preserved by pullbacks or products, *Proc. Camb. Philos. Soc.* 67 (1970), 553–558; MR 40/8024.
[277] B. J. Day and R. H. Street, Localisation of locally presentable categories, *J. Pure Appl. Alg.* 58 (1989), 227–233; MR 90h:18012.
[278] B. J. Day and R. H. Street, Localisations of locally presentable categories, II, *J. Pure Appl. Alg.* 63 (1990), 225–229; MR 91h:18007.
[279] J. P. Delale, Ensemble sous-jacent dans un topos, *C. R. Acad. Sci. Paris* 277 (1973), A153–156; MR 51/8197.
[280] R. Diaconescu, Change of base for toposes with generators, *J. Pure Appl. Alg.* 6 (1975), 191–218; MR 52/532.
[281] R. Diaconescu, Axiom of choice and complementation, *Proc. Amer. Math. Soc.* 51 (1975), 176–178; MR 51/10093.
[282] R. Diaconescu, Grothendieck toposes have Boolean points — a new proof, *Commun. Alg.* 4 (1976), 723–729; MR 54/2757.
[283] R. Diaconescu, On comparability in a topos, *Proc. Amer. Math. Soc.* 98 (1986), 389–393; MR 87m:18006.
[284] Y. Diers, Familles universelles de morphismes, *Ann. Soc. Sci. Bruxelles* 93 (1979), 175–195; MR 81i:18001.
[285] Y. Diers, Multimonads and multimonadic categories, *J. Pure Appl. Alg.* 17 (1980), 153–170; MR 81f:18024.
[286] Y. Diers, Catégories multialgébriques, *Arch. Math. (Basel)* 34 (1980), 193–209; MR 82m:18003a.
[287] Y. Diers, Catégories localement multiprésentables, *Arch. Math. (Basel)* 34 (1980), 344–356; MR 82m:18003b.
[288] Y. Diers, Un critère de représentabilité par sections continues de faisceaux, in *Category theory*, Lecture Notes in Math. vol. 962 (Springer–Verlag, 1982), 51–61; MR 82h:18006.

[289] Y. Diers, Une déscription axiomatique des catégories de faisceaux de structures algébriques sur les espaces topologiques booléens, *Adv. Math.* 47 (1983), 258–299; MR 84e:18007.

[290] Y. Diers, *Categories of Boolean sheaves of simple algebras*, Lecture Notes in Math. vol. 1187 (Springer–Verlag, 1986); MR 87j:18006.

[291] S. Di Sieno and M. Galuzzi, Categorie basate su un'algebra di Heyting, *Atti Accad. Sci. Torino* (I) 112 (1978), 333–343; MR 80h:18005.

[292] H. Dobbertin, Measurable refinement monoids and applications to distributive semilattices, Heyting algebras, and Stone spaces, *Math. Z.* 187 (1984), 13–21; MR 85h:20072.

[293] D. E. Dobbs, On characterizing injective sheaves, *Canad. J. Math.* 29 (1977), 1031–1039; MR 56/5685; corrigenda *ibid.* 32 (1980), 1522; MR 82f:18015b.

[294] L. van den Dries, Some applications of a model-theoretic fact to (semi-)algebraic geometry, *Proc. Kon. Nederl. Akad. Weten.* A 85 (1982), 397–401 = *Indag. Math.* 44 (1982), 397–401; MR 84m:14029.

[295] I. de F. Druck, Un modèle de filtres pour l'analyse réelle synthétique, *Cahiers Top. Géom. Diff. Catégoriques* 34 (1993), 83–120; MR 94g:18001.

[296] E. J. Dubuc, Free monoids, *J. Algebra* 29 (1974), 208–228; MR 50/13191.

[297] E. J. Dubuc, Concrete quasitopoi, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 239–254; MR 81e:18010.

[298] E. J. Dubuc, Sur les modèles de la géométrie différentielle synthétique, *Cahiers Top. Géom. Diff.* 20 (1979), 234–279; MR 81c:58003.

[299] E. J. Dubuc, C^∞-schemes, *Amer. J. Math.* 103 (1981), 683–690; MR 83a:58004.

[300] E. J. Dubuc, Open covers and infinitary operations in C^∞-rings, *Cahiers Top. Géom. Diff.* 22 (1981), 287–300; MR 83m:18010.

[301] E. J. Dubuc, Logical opens and real numbers in topoi, *J. Pure Appl. Alg.* 43 (1986), 129–143; MR 88d:18003.

[302] E. J. Dubuc, Integration of vector fields and synthetic differential geometry (Spanish), *Rev. Un. Mat. Argentina* 35 (1989), 151–162; MR 95c:58007.

[303] E. J. Dubuc, Germ representability and local integration of vector fields in a well adapted model of SDG, *J. Pure Appl. Alg.* 64 (1990), 131–144; MR 91g:18011.

[304] E. J. Dubuc, Topoi and (multisorted) higher-order intuitionistic logic, in *Séminaire sur les structures algébriques ordonnées*, vol. II, Publ. Math. Univ. Paris VII no. 33 (1990), 127–140; MR 91h:03090.

[305] E. J. Dubuc, Axiomatic etal maps and a theory of spectrum, *J. Pure Appl. Alg.* 149 (2000), 15–45; MR 02a:18006.

[306] E. J. Dubuc and G. M. Kelly, A presentation of topoi as algebraic relative to categories or graphs, *J. Algebra* 81 (1983), 420–433; MR 85c:18009.

[307] E. J. Dubuc and A. Kock, On 1-form classifiers, *Commun. Alg.* 12 (1984), 1471–1531; MR 86b:58007.

[308] E. J. Dubuc and J. Penon, Objets compacts dans les topos, *J. Austral. Math. Soc.* (A) 40 (1986), 203–217; MR 87k:18004.

[309] E. J. Dubuc and G. E. Reyes, Subtoposes of the ring classifier, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 101–122; MR 81j:03098.

[310] E. J. Dubuc and G. Taubin, Analytic rings, *Cahiers Top. Géom. Diff.* 24 (1983), 225–265; MR 86j:32011.

[311] E. J. Dubuc and J. G. Zilber, On analytic models of synthetic differential geometry, *Cahiers Top. Géom. Diff. Catégoriques* 35 (1994), 49–73; MR 95f:18007.

[312] E. J. Dubuc and J. G. Zilber, Banach spaces in an analytic model of synthetic differential geometry, *Cahiers Top. Géom. Diff. Catégoriques* 39 (1998), 117–136; MR 99h:58003.

[313] E. J. Dubuc and J. G. Zilber, Infinitesimal and local structures for Banach spaces and its exponentials in a topos, *Cahiers Top. Géom. Diff. Catégoriques* 41 (2000), 82–100.
[314] E. J. Dubuc and J. G. Zilber, Inverse function theorems for Banach spaces in a topos, *Cahiers Top. Géom. Diff. Catégoriques* 41 (2000), 207–224; MR 01f:18006.
[315] P. Dupont, Projectivity of sheaves on a locale and of free modules, *J. Pure Appl. Alg.* 62 (1989), 7–18; MR 90k:18003.
[316] P. Dupont, Rings and modules in a localic topos, *Bull. Soc. Math. Belg.* (B) 41 (1989), 63–71; MR 91c:18002.
[317] P. Dupont and B. Loiseau, A logical glance at sheaves, *Ann. Soc. Sci. Bruxelles* (I) 101 (1987), 133–149; MR 90m:03108.
[318] J. W. Duskin, *Simplicial methods and the interpretation of 'triple' cohomology*, Memoirs Amer. Math. Soc. no. 163 (1975); MR 52/14006.
[319] J. W. Duskin, Higher-dimensional torsors and the cohomology of topoi: the abelian theory, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 255–279; MR 82c:18006.
[320] J. W. Duskin, An outline of non-abelian cohomology in a topos, I: the theory of bouquets and gerbes, *Cahiers Top. Géom. Diff.* 23 (1982), 165–191; MR 83k:14020.
[321] R. Dyckhoff and W. Tholen, Exponentiable morphisms, partial products and pullback complements, *J. Pure Appl. Alg.* 49 (1987), 103–116; MR 88m:18003.
[322] M. M. Ebrahimi, Algebra in a Grothendieck topos: injectivity in quasi-equational classes, *J. Pure Appl. Alg.* 26 (1982), 269–280; MR 84d:18010.
[323] M. M. Ebrahimi, Equational compactness of sheaves of algebras on a Noetherian locale, *Alg. Universalis* 16 (1983), 318–330; MR 84e:08010.
[324] M. M. Ebrahimi, Bimorphisms and tensor products in a topos, *Bull. Iran. Math. Soc.* 15 (1988), 1–31; MR 91g:08002.
[325] M. M. Ebrahimi, Internal completeness and injectivity of Boolean algebras in the topos of M-sets, *Bull. Austral. Math. Soc.* 41 (1990), 323–332; MR 91b:03102.
[326] M. M. Ebrahimi and M. Mahmoudi, When is the category of separated M-sets a quasitopos or a topos?, *Bull. Iran. Math. Soc.* 21 (1995), 25–33; MR 98a:18004.
[327] M. M. Ebrahimi and M. Mahmoudi, The category of M-sets, *Ital. J. Pure Appl. Math.* 9 (2001), 123–132.
[328] K. Eda, A minimal flabby sheaf and an abelian group, *Tsukuba Math. J.* 7 (1983), 157–168; MR 85d:18007.
[329] K. R. Edwards, *Relative finiteness and the preservation of filtered colimits*, Ph.D. thesis, University of Chicago (1980).
[330] C. Ehresmann, Gattungen von lokalen Strukturen, *Jber. Deutsch. Math.-Verein.* 60 (1957), 59–77; MR 20/2392.
[331] C. Ehresmann, *Catégories et structures* (Dunod, 1965); MR 30/4809.
[332] C. Ehresmann, Esquisses et types de structures algébriques, *Bul. Inst. Politehn. Iaşi* 14 (1968), 1–14; MR 39/278.
[333] T. Ehrhard, Dictoses, in *Category theory and computer science*, Lecture Notes in Computer Science vol. 389 (Springer–Verlag, 1989), 213–223; MR 91b:68070.
[334] D. P. Ellerman, Sheaves of structures and generalized ultraproducts, *Ann. Math. Logic* 7 (1974), 163–195; MR 51/12516.
[335] H. Engenes, Subobject classifiers and classes of subfunctors, *Math. Scand.* 34 (1974), 145–152; MR 51/8200.
[336] Yu.L. Ershov, Computable numerations of morphisms (Russian), *Algebra i Logika* 10 (1971), 247–308; English translation in *Algebra and Logic* 10 (1971), 155–191; MR 46/5124.

[337] L. Esakia, M. Jibladze and D. Pataraia, Scattered toposes, *Ann. Pure Appl. Logic* 103 (2000), 97–107; MR 01g:03121.

[338] M. H. Escardó, The regular-locally-compact reflection of a stably locally compact locale, *J. Pure Appl. Alg.* 157 (2001), 41–55.

[339] L. Español, Le spectre d'un anneau dans l'algèbre constructive et applications à la dimension, *Cahiers Top. Géom. Diff.* 24 (1983), 133–144; MR 85e:13034.

[340] L. Español, A note on integral domains in toposes (Spanish), in *Proc. Tenth Spanish–Portuguese Conf. on Math.*, vol. I (Univ. Murcia, 1985), 142–147; MR 87j:18007.

[341] L. Español, The spectrum and the ring dimension in algebra over a topos (Spanish), *Rev. Colomb. Mat.* 20 (1986), 105–128; MR 89g:18007.

[342] L. Español, The spectrum lattice of Baer rings and polynomials, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 118–124; MR 90e:18002.

[343] L. Español and L. Lambán, A tensor–hom adjunction in a topos related to vector topologies and bornologies, *J. Pure Appl. Alg.* 154 (2000), 143–158; MR 01f:18007.

[344] L. Español and M. C. Minguez, Logical open sets of convenient spaces in the Cahiers topos, in *Proc. Twelfth Portuguese-Spanish Conf. on Math.*, vol. II (Univ. do Minho, 1987), 597–602.

[345] M. Eytan, Engeler algorithms in a topos, in *Fundamentals of computation theory*, Math. Research vol. 2 (Akademie–Verlag, 1979), 128–137; MR 81g:03081.

[346] M. Eytan, Fuzzy sets: a topos-logical point of view, *Fuzzy Sets Syst.* 5 (1981), 47–67; MR 82k:03087.

[347] S. Fakir, Monade idempotente associée à une monade, *C. R. Acad. Sci. Paris* 270 (1970), A99–101; MR 41/1828.

[348] M. F. Farina, Una costruzione del fascio associato, *Ist Lombardo Accad. Sci. Lett. Rend.* (A) 110 (1976), 332–341; MR 58/16841.

[349] M. F. Farina and G. C. Meloni, Il funtore fascio associato nei topoi: la costruzione X^+, *Boll. Un. Mat. Ital.* A (5) 15 (1978), 639–647; MR 80b:18006.

[350] M. F. Farina and G. C. Meloni, The associated sheaf functor, *Ist. Lombardo Accad. Sci. Lett. Rend.* (A) 114 (1980), 50–55; MR 85d:18008.

[351] B. Fawcett and R. J. Wood, Constructive complete distributivity, I, *Math. Proc. Camb. Philos. Soc.* 107 (1990), 81–89; MR 90k:06017.

[352] P. Feit, *Axiomization of passage from 'local' structure to 'global' object*, Memoirs Amer. Math. Soc. no. 485 (1993); MR 94f:18009.

[353] Y. Félix and R. Lavendhomme, On De Rham's theorem in synthetic differential geometry, *J. Pure Appl. Alg.* 69 (1991), 21–31; MR 92a:58003.

[354] M. P. Fiore and G. D. Plotkin, An extension of models of axiomatic domain theory to models of synthetic domain theory, in *Computer science logic (Utrecht 1996)*, Lecture Notes in Computer Science vol. 1258 (Springer–Verlag, 1997), 129–139; MR 98k:03124.

[355] M. P. Fiore and G. Rosolini, Two models of synthetic domain theory, *J. Pure Appl. Alg.* 116 (1997), 151–162; MR 98j:18003.

[356] M. Fontana and G. Mazzola, Arithmétique fonctorielle des topologies de Grothendieck, *Ann. Univ. Ferrara Sez. VII* (N. S.) 22 (1976), 49–94; MR 56/12097.

[357] M. P. Fourman, Comparaison des réelles d'un topos; structures lisses sur un topos élémentaire, *Cahiers Top. Géom. Diff.* 16 (1976), 233–239.

[358] M. P. Fourman, The logic of topoi, in *Handbook of mathematical logic*, Studies in Logic and the Foundations of Math. vol. 90 (North–Holland, 1978), 1053–1090; MR 58/10343.

[359] M. P. Fourman, Sheaf models for set theory, *J. Pure Appl. Alg.* 19 (1980), 91–101; MR 82i:03078.

[360] M. P. Fourman, Notions of choice sequence, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 91–105; MR 85d:03115.

[361] M. P. Fourman, T_1 spaces over topological sites, *J. Pure Appl. Alg.* 27 (1983), 223–224; MR 84m:18004.

[362] M. P. Fourman, Continuous truth I: non-constructive objects, in *Logic Colloquium '82*, Studies in Logic and the Foundations of Math. vol. 112 (North–Holland, 1984), 161–180; MR 86b:03084.

[363] M. P. Fourman and R. J. Grayson, Formal spaces, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 107–122; MR 85c:03023.

[364] M. P. Fourman and J. M. E. Hyland, Sheaf models for analysis, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 280–301; MR 82g:03069.

[365] M. P. Fourman and A. Ščedrov, The 'world's simplest axiom of choice' fails, *Manuscripta Math.* 38 (1982), 325–332; MR 83k:03074.

[366] M. P. Fourman and D. S. Scott, Sheaves and logic, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 302–401; MR 82d:03061.

[367] R. H. Fox, On topologies for function-spaces, *Bull. Amer. Math. Soc.* 51 (1945), 429–432; MR 6–278.

[368] T. Fox, Coalgebras and cartesian categories, *Commun. Alg.* 4 (1976), 665–667; MR 53/8189.

[369] M. Fréchet, Sur quelques points du calcul fonctionnel, *Rend. Circ. Mat. Palermo* 22 (1906), 1–71.

[370] P. J. Freyd, The theories of functors and models, in *The theory of models*, Studies in Logic and the Foundations of Math. (North–Holland, 1965), 107–120; MR 34/1371.

[371] P. J. Freyd, Aspects of topoi, *Bull. Austral. Math. Soc.* 7 (1972), 1–76; MR 53/576; corrigenda *ibid.*, 467–480; MR 54/7571.

[372] P. J. Freyd, The axiom of choice, *J. Pure Appl. Alg.* 19 (1980), 103–125; MR 82i:03079.

[373] P. J. Freyd, Numerology in topoi, preprint, c. 1980.

[374] P. J. Freyd, All topoi are localic; or, Why permutation models prevail, *J. Pure Appl. Alg.* 46 (1987), 49–58; MR 88e:03091.

[375] P. J. Freyd, Choice and well-ordering, *Ann. Pure Appl. Logic* 35 (1987), 149–166; MR 88g:03093.

[376] P. J. Freyd, Combinators, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 63–66; MR 90m:03029.

[377] P. J. Freyd, H. Friedman and A. Scedrov, Lindenbaum algebras of intuitionistic theories and free categories, *Ann. Pure Appl. Logic* 35 (1987), 167–172; MR 88g:03094.

[378] P. J. Freyd, P. S. Mulry, G. Rosolini and D. S. Scott, Extensional PERs, in *Proc. Fifth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1990), 346–354; also published in *Info. Comput.* 98 (1992), 211–227; MR 93f:68065.

[379] P. J. Freyd, E. P. Robinson and G. Rosolini, Dinaturality for free, in *Applications of categories in computer science*, L.M.S. Lecture Notes Series no. 177 (Cambridge Univ. Press, 1992), 107–118; MR 94a:68084.

[380] P. J. Freyd, E. P. Robinson and G. Rosolini, Functorial parametricity, in *Proc. Seventh IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1992), 444–452.
[381] P. J. Freyd and A. Scedrov, *Categories, allegories* (North–Holland, 1990); MR 93c:18001.
[382] P. J. Freyd and R. H. Street, On the size of categories, *Theory Appl. Categ.* 1 (1995), 174–178; MR 96m:18003.
[383] J. Funk, The display locale of a cosheaf, *Cahiers Top. Géom. Diff. Catégoriques* 36 (1995), 53–93; MR 96d:18007.
[384] J. Funk, The locally connected coclosure of a Grothendieck topos, *J. Pure Appl. Alg.* 137 (1999), 17–27; MR 00b:18015.
[385] J. Funk, On branched covers in topos theory, *Theory Appl. Categ.* 7 (2000), 1–22; MR 00m:18008.
[386] P. Gabriel and F. Ulmer, *Lokal präsentierbare Kategorien*, Lecture Notes in Math. vol. 221 (Springer–Verlag, 1971); MR 48/6205.
[387] P. Gabriel and M. Zisman, *Calculus of fractions and homotopy theory*, Ergebnisse der Math. Band 35 (Springer–Verlag, 1967); MR 35/1019.
[388] F. Gago Couso, Compactness and Kock's principle (Galician), in *Proc. Twelfth Portuguese–Spanish Conf. on Math.*, vol. II (Univ. do Minho, 1987), 80–85.
[389] F. Gago Couso, Morse germs in synthetic differential geometry, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 125–129; MR 90a:58002.
[390] F. Gago Couso, Singularités dans la géométrie différentielle synthétique, *Bull. Soc. Math. Belg.* A 41 (1989), 279–287; MR 91b:18014.
[391] R. Gates, On the generic solution to $P(X) \cong X$ in distributive categories, *J. Pure Appl. Alg.* 125 (1998), 191–212; MR 99b:18013.
[392] R. Gates, On generic separable objects, *Theory Appl. Categ.* 4 (1998), 208–248; MR 00i:18015.
[393] N. Gavioli, $\mathcal{V}$-matrices over a topos, *Rend. Circ. Mat. Palermo Suppl.* (II) 24 (1990), 375–386; MR 92e:18004.
[394] M. Geanău, La catégorie des objets groupes abéliens dans un topos élémentaire est une catégorie abélienne, *Bul. Inst. Politehn. Bucureşti* 42 (1980), 9–19; MR 82c:18007.
[395] M. Geanău, Sur un invariant dans un topos élémentaire, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 27 (1981), 247–251; MR 83c:18003.
[396] M. Geanău, Objets injectifs dans la catégorie $\mathbf{ab}(\xi)$ des objets groupes abéliens dans un topos élémentaire, *Bul. Inst. Politehn. Bucureşti* 43 (1981), 9–14; MR 83i:18005.
[397] M. Geanău, Cohomologie abélienne dans un topos élémentaire, *Bull. Math. Soc. Sci. Math. Roum.* 26 (74) (1982), 51–57; MR 83g:18008.
[398] M. Geanău, L'orbite dans un topos élémentaire, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 30 (1984), 3–6; MR 87c:18001.
[399] M. Geanău, Quelques propriétés du topos des G-objets, *Bul. Inst. Politehn. Bucureşti* 46 (1985), 17–24; MR 87i:18003.
[400] S. Ghilardi, Presheaf semantics and independence results for some nonclassical first-order logics, *Arch. Math. Logic* 29 (1989), 125–136; MR 91f:03048.
[401] S. Ghilardi, Incompleteness results in Kripke semantics, *J. Symbolic Logic* 56 (1991), 517–538; MR 93b:03020.
[402] S. Ghilardi, Free Heyting algebras as bi-Heyting algebras, *C. R. Math. Rep. Acad. Sci. Canada* 24 (1992), 240–244.

[403] S. Ghilardi and M. W. Zawadowski, A sheaf representation and duality for finitely presented Heyting algebras, *J. Symbolic Logic* 60 (1995), 911–939; MR 96i:03063.

[404] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove and D. S. Scott, *A compendium of continuous lattices* (Springer–Verlag, 1980); MR 82h:06005.

[405] J. Giraud, Analysis situs, *Séminaire Bourbaki*, exposé 256 (1963); MR 33/1343.

[406] J. Giraud, *Méthode de la descente*, Bull. Soc. Math. France, Mémoire 2 (1964); MR 32/7556.

[407] J. Giraud, *Cohomologie non abélienne*, Grundl. Math. Wissen. Band 179 (Springer–Verlag, 1971); MR 49/8992.

[408] J. Giraud, Classifying topos, in *Toposes, algebraic geometry and logic*, Lecture Notes in Math. vol. 274 (Springer–Verlag, 1972), 43–56; MR 50/2300.

[409] A. M. Gleason, Projective topological spaces, *Ill. J. Math.* 2 (1958), 482–489; MR 22/12509.

[410] P. G. Goerss and J. F. Jardine, Localization theories for simplicial presheaves, *Canad. J. Math.* 50 (1998), 1048–1089; MR 99j:55009.

[411] R. I. Goldblatt, *Topoi: the categorial analysis of logic*, Studies in Logic and the Foundations of Math. vol. 98 (North–Holland, 1979; revised second edition, 1984); MR 81a:03063 and 85m:03002.

[412] R. I. Goldblatt, Grothendieck topology as geometric modality, *Z. Math. Logik Grundl. Math.* 27 (1981), 495–529; MR 83d:03069.

[413] I. Graham, Fuzzy sets and toposes — towards higher order fuzzy logic, *Fuzzy Sets Syst.* 23 (1987), 19–32; MR 88g:03071.

[414] J. W. Gray, Sheaves with values in a category, *Topology* 3 (1964), 1–18; MR 28/2545.

[415] J. W. Gray, Fragments of the history of sheaf theory, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 1–79; MR 82j:01060.

[416] R. J. Grayson, Heyting-valued models for intuitionistic set theory, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 402–414; MR 82g:03089.

[417] R. J. Grayson, Concepts of general topology in constructive mathematics and in sheaves, *Ann. Math. Logic* 20 (1981), 1–41; MR 82g:03105; corrigenda *ibid.* 23 (1982), 99; MR 84d:03075.

[418] R. J. Grayson, Concepts of general topology in constructive mathematics and in sheaves, II, *Ann. Math. Logic* 23 (1982), 55–98; MR 84c:03102.

[419] P. A. Grillet, Regular categories, in *Exact categories and categories of sheaves*, Lecture Notes in Math. vol. 236 (Springer–Verlag, 1971), 121–222.

[420] A. Grothendieck, Sur quelques points d'algèbre homologique, *Tohôku Math. J.* 9 (1957), 119–221; MR 21/1328.

[421] A. Grothendieck, Technique de descente et théorèmes d'existence en géométrie algébrique, I–VI, *Sém. Bourbaki*, exposés 190/195/212/221/232/236 (1959–62); MR 26/3566.

[422] A. Grothendieck, *Revêtements étales et groupe fondamental*, Séminaire de Géométrie Algébrique du Bois-Marie, année 1960–61; second edition published as Lecture Notes in Math. vol. 224 (Springer–Verlag, 1971); MR 50/7129.

[423] J. H. C. Gunawardena, Geometric logic, causality and event structures, in *CONCUR '91*, Lecture Notes in Computer Science vol. 527 (Springer–Verlag, 1991); MR 93b:03043.

[424] A. K. Guts, A topos-theoretic approach to the foundations of relativity theory (Russian), *Dokl. Akad. Nauk. SSSR* 318 (1991), 1294–1297; English translation in *Soviet Math. Dokl.* 43 (1991), 904–907; MR 92m:83083.

[425] J. Haigh, Essential geometric morphisms between toposes over finite sets, *Math. Proc. Camb. Philos. Soc.* 87 (1980), 21–24; MR 80k:18006.

[426] M. Hakim, *Topos annelés et schémas relatifs*, Ergebnisse der Math. Band 64 (Springer–Verlag, 1972); MR 51/500.
[427] A. W. Hales, On the non-existence of free complete Boolean algebras, *Fund. Math.* 54 (1964), 45–66; MR 29/1162.
[428] R. Harting, A remark on injectivity of sheaves of abelian groups, *Arch. Math. (Basel)* 36 (1981), 230–233; MR 82j:20101.
[429] R. Harting, Locally injective G-sheaves of abelian groups, *Cahiers Top. Géom. Diff.* 22 (1981), 115–122; MR 84a:18022.
[430] R. Harting, Abelian groups in an elementary topos, *Commun. Alg.* 10 (1982), 1173–1237; MR 83i:18006.
[431] R. Harting, Locally injective abelian groups in a topos, *Commun. Alg.* 11 (1983), 349–376; MR 84m:18005.
[432] R. Harting, Abelian groups in a topos: injectives and injective effacements, *J. Pure Appl. Alg.* 30 (1983), 247–260; MR 85h:18002.
[433] S. Hayashi, On set theories in toposes, in *Logic symposia, Hakone 1979–80*, Lecture Notes in Math. vol. 891 (Springer–Verlag, 1981), 23–29; MR 84b:18002.
[434] W. He and Y. M. Liu, Inverse limits in the category of locales, *Sci. China* (A) 41 (1998), 476–482; MR 99i:18008.
[435] W. He and Y. M. Liu, Steenrod's theorem for locales, *Math. Proc. Camb. Philos. Soc.* 124 (1998), 305–307; MR 99c:18001.
[436] M. Hébert, Characterizations of axiomatic categories of models canonically isomorphic to (quasi-) varieties, *Canad. Math. Bull.* 31 (1988), 287–300; MR 90d:03064.
[437] M. Hébert, Syntactic characterization of closure under connected limits, *Arch. Math. Logic* 31 (1991), 133–143; MR 93e:03050.
[438] M. Hébert, Syntactic characterization of closure under pullbacks and of locally polypresentable categories, *Ann. Pure Appl. Logic* 84 (1997), 73–95; MR 97k:18003.
[439] A. Heller and K. A. Rowe, On the category of sheaves, *Amer. J. Math.* 84 (1962), 205–216; MR 26/1887.
[440] D. Higgs, Boolean-valued equivalence relations and complete extensions of complete Boolean algebras, *Bull. Austral. Math. Soc.* 3 (1970), 65–72; MR 42/156.
[441] D. Higgs, A category approach to Boolean-valued set theory, Lecture notes, Univ. of Waterloo (1973).
[442] D. Higgs, A universal characterization of $[0,\infty]$, *Proc. Kon. Nederl. Akad. Weten.* A 81 (1978) = *Indag. Math.* 40 (1978), 448–455; MR 80a:08007.
[443] D. Higgs, Injectivity in the topos of complete Heyting algebra valued sets, *Canad. J. Math.* 36 (1984), 550–568; MR 85m:18003.
[444] H. L. Hiller, Fibrations and Grothendieck topologies, *Bull. Austral. Math. Soc.* 14 (1976), 111–128; MR 53/13355.
[445] M. Hochster, Prime ideal structure in commutative rings, *Trans. Amer. Math. Soc.* 142 (1969), 43–60; MR 40/4257.
[446] W. Hodges, *Model theory*, Encyclopaedia of Mathematics and its Applications vol. 42 (Cambridge Univ. Press, 1993); MR 94e:03002.
[447] G. van der Hoeven and I. Moerdijk, On an independence result in the theory of lawless sequences, *Proc. Kon. Nederl. Akad. Weten.* A 86 (1983) = *Indag. Math.* 45 (1983), 185–191; MR 85d:03116.
[448] G. van der Hoeven and I. Moerdijk, Sheaf models for choice sequences, *Ann. Pure Appl. Logic* 27 (1984), 63–107; MR 86m:03091.
[449] G. van der Hoeven and I. Moerdijk, On choice sequences determined by spreads, *J. Symbolic Logic* 49 (1984), 908–916; MR 87m:03093.

[450] K. H. Hofmann, Stably continuous frames, and their topological manifestations, in *Categorical topology*, Sigma Ser. Pure Math. no. 5 (Heldermann–Verlag, 1984), 282–307; MR 86h:06019.

[451] K. H. Hofmann and K. Keimel, Sheaf-theoretical concepts in analysis, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 415–441; MR 81f:46085.

[452] K. H. Hofmann and J. D. Lawson, The spectral theory of distributive continuous lattices, *Trans. Amer. Math. Soc.* 246 (1978), 285–310; MR 80c:54045.

[453] U. Höhle, Fuzzy topologies and topological space objects in a topos, *Fuzzy Sets Syst.* 19 (1986), 299–304; MR 87m:18007.

[454] U. Höhle, Monoidal closed categories, weak topoi, and generalized logics, *Fuzzy Sets Syst.* 42 (1991), 15–35; MR 92i:03071.

[455] U. Höhle and L. N. Stout, Foundations of fuzzy sets, *Fuzzy Sets Syst.* 40 (1991), 257–296; MR 92c:04009.

[456] D. A. J. Hoskin, The Stein topos, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 154–167; MR 85c:32027.

[457] D. Howe, Module categories over topoi, *J. Pure Appl. Alg.* 21 (1981), 161–165; MR 82m:18002.

[458] H. Hu and W. Tholen, A note on free regular and exact completions and their infinitary generalizations, *Theory Appl. Categ.* 2 (1996), 113–132; MR 98e:18001.

[459] V. Huber-Dyson, *Gödel's Theorems: a workbook on formalization*, Teubner-Texte zur Mathematik, Band 122 (Teubner–Verlag, 1991); MR 93a:03001.

[460] J. M. E. Hyland, Continuity in spatial toposes, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 442–465; MR 83g:18009.

[461] J. M. E. Hyland, Function spaces in the category of locales, in *Continuous lattices*, Lecture Notes in Math. vol. 871 (Springer–Verlag, 1981), 264–281.

[462] J. M. E. Hyland, The effective topos, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 165–216; MR 84m:03101.

[463] J. M. E. Hyland, A small complete category, *Ann. Pure Appl. Logic* 40 (1988), 135–165; MR 90a:18007.

[464] J. M. E. Hyland, First steps in synthetic domain theory, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 131–156; MR 93g:68075.

[465] J. M. E. Hyland, P. T. Johnstone and A. M. Pitts, Tripos theory, *Math. Proc. Camb. Philos. Soc.* 88 (1980), 205–232; MR 81i:03102.

[466] J. M. E. Hyland and C. H. L. Ong, Modified realizability toposes and strong normalization proofs, in *Typed lambda calculi and applications*, Lecture Notes in Computer Science vol. 664 (Springer–Verlag, 1993), 179–194.

[467] J. M. E. Hyland and A. M. Pitts, The theory of constructions: categorical semantics and topos-theoretic models, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 137–199; MR 90h:03023.

[468] J. M. E. Hyland, E. P. Robinson and G. Rosolini, The discrete objects in the effective topos, *Proc. London Math. Soc.* (3) 60 (1990), 1–36; MR 90k:18004.

[469] J. M. E. Hyland, E. P. Robinson and G. Rosolini, Algebraic types in PER models, in *Mathematical foundations of programming semantics*, Lecture Notes in Computer Science vol. 442 (Springer–Verlag, 1990), 333–350; MR 91h:68128.

[470] P. M. Idziak, Sheaves in universal algebra and model theory, I, *Rep. Math. Logic* 23 (1989), 39–65; MR 92g:08010a.

[471] P. M. Idziak, Sheaves in universal algebra and model theory, II, *Rep. Math. Logic* 24 (1990), 61–86; MR 92g:08010b.

[472] C. K. Im, On elementary toposes, *Bull. Korean Math. Soc.* 16 (1980), 55–71; MR 81h:18006.

[473] C. Ionită, Characterization of Lawvere regular categories with exact filtered colimits and generators, *Rev. Roum. Math Pures Appl.* 27 (1981), 43–59; MR 83g:18010.

[474] C. Ionită, Algebraic toposes and topologies, *Rev. Roum. Math Pures Appl.* 27 (1982), 831–847; MR 84g:18013.

[475] J. R. Isbell, General functorial semantics, I, *Amer. J. Math.* 94 (1972), 535–596; MR 53/580.

[476] J. R. Isbell, Atomless parts of spaces, *Math. Scand.* 31 (1972), 5–32; MR 50/11184.

[477] J. R. Isbell, Function spaces and adjoints, *Math. Scand.* 36 (1975), 317–339; MR 53/9134.

[478] J. R. Isbell, Product spaces in locales, *Proc. Amer. Math. Soc.* 81 (1981), 116–118; MR 82c:54006.

[479] J. R. Isbell, Direct limits of meet-continuous lattices, *J. Pure Appl. Alg.* 23 (1982), 33–35; MR 83a:18015.

[480] J. R. Isbell, Completion of a construction of Johnstone, *Proc. Amer. Math. Soc.* 85 (1982), 333–334; MR 83i:06011.

[481] J. R. Isbell, Varieties of frames, *Bull. London Math. Soc.* 17 (1985), 41–42; MR 86c:08011.

[482] J. R. Isbell, Graduation and dimension in locales, in *Aspects of topology in memory of Hugh Dowker*, L.M.S. Lecture Notes Series no. 93 (Cambridge Univ. Press, 1985), 195–210; MR 86h:54042.

[483] J. R. Isbell, General function spaces, products and continuous lattices, *Math. Proc. Camb. Philos. Soc.* 100 (1986), 193–205; MR 87k:54023.

[484] J. R. Isbell, First steps in descriptive theory of locales, *Trans. Amer. Math. Soc.* 327 (1991), 353–371; MR 92b:54078; corrigenda *ibid.* 341 (1994), 467–468; MR 94c:54071.

[485] J. R. Isbell, On dissolute spaces, *Top. Appl.* 40 (1991), 63–70; MR 92d:18002.

[486] J. R. Isbell, I. Kříž, A. Pultr and J. Rosický, Remarks on localic groups, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 154–172; MR 90f:18009.

[487] J. R. Isbell and S. H. Schanuel, Weak powers and small-valued functors, *Bull. London Math. Soc.* 14 (1982), 221–222; MR 83h:18003.

[488] C. J. Isham, Topos theory and consistent histories: the internal logic of the set of all consistent sets, *Internat. J. Theor. Phys.* 36 (1997), 785–814; MR 99a:81014.

[489] C. J. Isham and J. Butterfield, Some possible roles for topos theory in quantum theory and quantum gravity, *Found. Phys.* 30 (2000), 1707–1735; MR 02c:83037.

[490] B. Jacobs, *Categorical logic and type theory*, Studies in Logic and the Foundations of Math. vol. 141 (North-Holland, 1999); MR 01b:03077.

[491] W. James, Sheaf spaces on finite closed sets, *Logique & Analyse* 137–8 (1992), 175–188; MR 98e:03041.

[492] A. W. Jankowski and M. W. Zawadowski, Sheaves over Heyting lattices, *Studia Logica* 44 (1985), 237–256; MR 88a:18004.

[493] P. Jara, A. Verschoren and C. Vidal, *Localization and sheaves: a relative point of view*, Pitman Research Notes in Math. vol. 339 (Longman, 1995); MR 97m:18011.

[494] J. F. Jardine, Simplicial objects in a Grothendieck topos, in *Applications of K-theory in algebraic geometry and number theory*, Contemp. Math. vol. 55 (Amer. Math. Soc., 1986), 193–239; MR 88g:18008.

[495] J. F. Jardine, Simplicial presheaves, *J. Pure Appl. Alg.* 47 (1987), 35–87; MR 88j:18005.

[496] C. B. Jay, Finite objects in a locos, *J. Pure Appl. Alg.* 116 (1997), 169–183; MR 98a:18003.

[497] M. Jibladze, Geometric morphisms and indexed toposes (Russian), *Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk. Gruzin. SSR* 91 (1988), 12–18; MR 90m:18005; English translation in *Categorical topology* (World Scientific, 1989), 10–18; MR 91c:18003.

[498] M. Jibladze, A presentation of the initial lift-algebra, *J. Pure Appl. Alg.* 116 (1997), 185–198; MR 98c:18002.

[499] M. Jibladze, Lower bagdomain as a glueing, *Proc. A. Razmadze Math. Inst.* 118 (1998), 33–41; MR 00b:18005.

[500] M. Jibladze and P. T. Johnstone, The frame of fibrewise closed nuclei, *Cahiers Top. Géom. Diff. Catégoriques* 32 (1991), 99–112; MR 93b:18010.

[501] P. T. Johnstone, The associated sheaf functor in an elementary topos, *J. Pure Appl. Alg.* 4 (1974), 231–242; MR 50/10002.

[502] P. T. Johnstone, Internal categories and classification theorems, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 103–113; MR 53/581.

[503] P. T. Johnstone, Adjoint lifting theorems for categories of algebras, *Bull. London Math. Soc.* 7 (1975), 294–297; MR 52/10845.

[504] P. T. Johnstone, *Topos theory*, L.M.S. Monographs no. 10 (Academic Press, 1977); MR 57/9791.

[505] P. T. Johnstone, Rings, fields and spectra, *J. Algebra* 49 (1977), 238–260; MR 56/8335.

[506] P. T. Johnstone, Automorphisms of Ω, *Alg. Universalis* 9 (1979), 1–7; MR 80f:18003.

[507] P. T. Johnstone, On a topological topos, *Proc. London Math. Soc.* (3) 38 (1979), 237–271; MR 80i:18010.

[508] P. T. Johnstone, A syntactic approach to Diers' localizable categories, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 466–478; MR 82k:03102.

[509] P. T. Johnstone, Conditions related to De Morgan's law, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 479–491; MR 81e:03062.

[510] P. T. Johnstone, Another condition equivalent to De Morgan's law, *Commun. Alg.* 7 (1979), 1309–1312; MR 81d:03067.

[511] P. T. Johnstone, Remarks on Kock's paper 'Formally real local rings, and infinitesimal stability', in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 137–138; MR 81b:03064.

[512] P. T. Johnstone, The Gleason cover of a topos, I, *J. Pure Appl. Alg.* 19 (1980), 171–192; MR 82a:18002.

[513] P. T. Johnstone, Open maps of toposes, *Manuscripta Math.* 31 (1980), 217–247; MR 81f:18020.

[514] P. T. Johnstone, Tychonoff's theorem without the axiom of choice, *Fund. Math.* 113 (1981), 21–35; MR 84h:03117.

[515] P. T. Johnstone, The Gleason cover of a topos, II, *J. Pure Appl. Alg.* 22 (1981), 229–247; MR 83a:18011.

[516] P. T. Johnstone, Scott is not always sober, in *Continuous lattices*, Lecture Notes in Math. vol. 871 (Springer–Verlag, 1981), 282–283.

[517] P. T. Johnstone, Injective toposes, in *Continuous lattices*, Lecture Notes in Math. vol. 871 (Springer–Verlag, 1981), 284–297.

[518] P. T. Johnstone, Factorization theorems for geometric morphisms, I, *Cahiers Top. Géom. Diff.* 22 (1981), 3–17; MR 82j:18007.

[519] P. T. Johnstone, Factorization theorems for geometric morphisms, II, in *Categorical aspects of topology and analysis*, Lecture Notes in Math. vol. 915 (Springer–Verlag, 1982), 216–233; MR 83j:18006.

[520] P. T. Johnstone, *Stone spaces*, Cambridge Studies in Advanced Math. no. 3 (Cambridge Univ. Press, 1982); MR 85f:54002.

[521] P. T. Johnstone, Quotients of decidable objects in a topos, *Math. Proc. Camb. Philos. Soc.* 93 (1983), 409–419; MR 84h:18006.

[522] P. T. Johnstone, Almost maximal ideals, *Fund. Math.* 123 (1984), 197–209; MR 86f:03081.

[523] P. T. Johnstone, Open locales and exponentiation, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 84–116; MR 85i:18005.

[524] P. T. Johnstone, How general is a generalized space?, in *Aspects of topology in memory of Hugh Dowker*, L.M.S. Lecture Notes Series no. 93 (Cambridge Univ. Press, 1985), 77–111; MR 87f:03173.

[525] P. T. Johnstone, When is a variety a topos?, *Alg. Universalis* 21 (1985), 198–212; MR 87j:18008.

[526] P. T. Johnstone, Vietoris locales and localic semilattices, in *Continuous lattices and their applications*, Lecture Notes Pure Appl. Math. vol. 101 (Marcel Dekker, 1985), 155–180; MR 87e:54026.

[527] P. T. Johnstone, A simple proof that localic subgroups are closed, *Cahiers Top. Géom. Diff. Catégoriques* 29 (1988), 157–161; MR 89e:18019.

[528] P. T. Johnstone, A constructive 'closed subgroup theorem' for localic groups and groupoids, *Cahiers Top. Géom. Diff. Catégoriques* 30 (1989), 3–23; MR 90e:22006.

[529] P. T. Johnstone, A topos-theorist looks at dilators, *J. Pure Appl. Alg.* 58 (1989), 235–249; MR 90f:18002.

[530] P. T. Johnstone, Collapsed topses as bitopological spaces, in *Categorical topology* (World Scientific, 1989), 19–35; MR 91g:18002.

[531] P. T. Johnstone, Collapsed toposes and cartesian closed varieties, *J. Algebra* 129 (1990), 446–480; MR 91c:18007.

[532] P. T. Johnstone, Two notes on nuclei, *Order* 7 (1990), 205–210; MR 92d:06028.

[533] P. T. Johnstone, Fibrewise separation axioms for locales, *Math. Proc. Camb. Philos. Soc.* 108 (1990), 247–256; MR 91h:06024.

[534] P. T. Johnstone, The 'closed subgroup theorem' for localic herds and pregroupoids, *J. Pure Appl. Alg.* 70 (1991), 97–106; MR 92c:22008.

[535] P. T. Johnstone, Partial products, bagdomains and hyperlocal toposes, in *Applications of categories in computer science*, L.M.S. Lecture Notes Series no. 177 (Cambridge Univ. Press, 1992), 315–339.

[536] P. T. Johnstone, Fibrations and partial products in a 2-category, *Appl. Categ. Struct.* 1 (1993), 141–179; MR 94i:18007.

[537] P. T. Johnstone, Variations on the bagdomain theme, *Theoret. Comput. Sci.* 136 (1994), 3–20; MR 96g:68070.

[538] P. T. Johnstone, Remarks on quintessential and persistent localizations, *Theory Appl. Categ.* 2 (1996), 90–99; MR 97j:18004.

[539] P. T. Johnstone, Cartesian monads on toposes, *J. Pure Appl. Alg.* 116 (1997), 199–220; MR 97m:18007.

[540] P. T. Johnstone, A note on discrete Conduché fibrations, *Theory Appl. Categ.* 5 (1999), 1–11; MR 00c:18001.
[541] P. T. Johnstone, An 'unsitely' result on atomic morphisms, *Appl. Categ. Struct.* 8 (2000), 7–15; MR 01d:18005.
[542] P. T. Johnstone and A. Joyal, Continuous categories and exponentiable toposes, *J. Pure Appl. Alg.* 25 (1982), 255–296; MR 83k:18005.
[543] P. T. Johnstone and F. E. J. Linton, Finiteness and decidability, II, *Math. Proc. Camb. Philos. Soc.* 84 (1978), 207–218; MR 81f:18016b.
[544] P. T. Johnstone, F. E. J. Linton and R. Paré, Injectives in topoi, II: connections with the axiom of choice, in *Categorical topology*, Lecture Notes in Math. vol. 719 (Springer–Verlag, 1979), 207–216; MR 80j:18007b.
[545] P. T. Johnstone and I. Moerdijk, Local maps of toposes, *Proc. London Math. Soc.* (3) 58 (1989), 281–305; MR 90a:18004.
[546] P. T. Johnstone, A. J. Power, T. Tsujishita, H. Watanabe and J. Worrell, The structure of categories of coalgebras, *Theoret. Comput. Sci.* 260 (2001), 87–117.
[547] P. T. Johnstone and E. P. Robinson, A note on inequivalence of realizability toposes, *Math. Proc. Camb. Philos. Soc.* 105 (1989), 1–3; MR 90b:03095.
[548] P. T. Johnstone and S. J. Vickers, Preframe presentations present, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 193–212; MR 93g:06006.
[549] P. T. Johnstone and G. C. Wraith, Algebraic theories in toposes, in *Indexed categories and their applications*, Lecture Notes in Math. vol. 661 (Springer–Verlag, 1978), 141–242; MR 81f:18019.
[550] D. de Jongh, A class of intuitionistic connectives, in *The Kleene symposium*, Studies in Logic and the Foundations of Math. vol. 101 (North–Holland, 1980), 103–111; MR 81m:03013.
[551] B. Jónsson and A. Tarski, On two properties of free algebras, *Math. Scand.* 9 (1961), 95–101; MR 23/A3695.
[552] A. Joyal, Une théorie combinatoire des séries formelles, *Adv. Math.* 42 (1981), 1–82; MR 84d:05025.
[553] A. Joyal and I. Moerdijk, Toposes are cohomologically equivalent to spaces, *Amer. J. Math.* 112 (1990), 87–95; MR 90m:18006.
[554] A. Joyal and I. Moerdijk, Toposes as homotopy groupoids, *Adv. Math.* 80 (1990), 22–38; MR 91b:18004.
[555] A. Joyal and I. Moerdijk, A categorical theory of cumulative hierarchies of sets, *C. R. Math. Rep. Acad. Sci. Canada* 13 (1991), 55–58; MR 92c:18002.
[556] A. Joyal and I. Moerdijk, A completeness theorem for open maps, *Ann. Pure Appl. Logic* 70 (1994), 51–86; MR 95j:03104; abstracted in *C. R. Math. Rep. Acad. Sci. Canada* 12 (1990), 253–256; MR 91m:18008.
[557] A. Joyal and I. Moerdijk, *Algebraic set theory*, L.M.S. Lecture Notes Series no. 220 (Cambridge Univ. Press, 1995); MR 96k:03127.
[558] A. Joyal and G. E. Reyes, Separably real closed local rings, *Rev. Colomb. Mat.* 19 (1985), 107–115; MR 88d:18004; also in *J. Pure Appl. Alg.* 43 (1986), 271–279; MR 89e:03054.
[559] A. Joyal and R. H. Street, Pullbacks equivalent to pseudopullbacks, *Cahiers Top. Géom. Diff. Catégoriques* 34 (1993), 153–156; MR 94a:18004.
[560] A. Joyal and M. Tierney, *An extension of the Galois theory of Grothendieck*, Memoirs Amer. Math. Soc. no. 309 (1984); MR 86d:18002.
[561] A. Joyal and M. Tierney, Universal bundles in simplicial sheaves, *C. R. Math. Rep. Acad. Sci. Canada* 12 (1990), 167–172; MR 91j:18020.
[562] A. Joyal and M. Tierney, Strong stacks and classifying spaces, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 213–236; MR 92c:18013; abstracted in *C. R. Math. Rep. Acad. Sci. Canada* 13 (1991), 2–6; MR 93h:18019.

[563] A. Joyal and M. Tierney, Classifying spaces for sheaves of simplicial groupoids, *J. Pure Appl. Alg.* 89 (1993), 135–161; MR 94i:18005.

[564] A. Joyal and M. Tierney, On the homotopy theory of sheaves of simplicial groupoids, *Math. Proc. Camb. Philos. Soc.* 120 (1996), 263–290; MR 97g:55025.

[565] A. Joyal and M. Tierney, On the theory of path groupoids, *J. Pure Appl. Alg.* 149 (2000), 69–100; MR 01j:55012.

[566] A. Joyal and G. C. Wraith, Eilenberg–Mac Lane toposes and cohomology, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 117–131; MR 85k:18006.

[567] P. Karazeris, Compact topologies on locally presentable categories, *Cahiers Top. Géom. Diff. Catégoriques* 38 (1997), 227–255; MR 98e:18003.

[568] P. Karazeris, Gabriel topologies on coherent quantales, *J. Pure Appl. Alg.* 127 (1998), 177–192; MR 99e:18004.

[569] S. Kasangian and S. Vigna, The topos of trees: a categorical semantics for SCCS, *Fund. Inform.* 32 (1997), 27–45; MR 99f:18007.

[570] Y. Katsov, Model theory of functors: axiomatizability problems, *J. Pure Appl. Alg.* 119 (1997), 285–296; MR 98m:03074.

[571] E. B. Katsov and E. G. Shul'geifer, On the groupoid of varieties of group obects of a topos of presheaves (Russian), in *Algebraic systems*, Volgograd Gos. Ped. Inst. (1989), 81–93; MR 94c:18014.

[572] Y. Kawahara, Pushout-complements and basic concepts of grammars in toposes, *Theoret. Comput. Sci.* 77 (1990), 267–289; MR 91m:68113.

[573] Y. Kawahara and Y. Mizoguchi, Categorical assertion semantics in topoi, *Adv. Software Sci. Tech.* 4 (1992), 137–149.

[574] A. Kayihura and M. M. Mawanda, On a categorical analysis of Zadeh generalized subsets of sets, II, *Fuzzy Sets Syst.* 51 (1992), 219–226; MR 93m:18006.

[575] O. Keane, Abstract Horn theories, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 15–50; MR 52/2871.

[576] W. F. Keigher, Differential algebra in a topos, *Cahiers Top. Géom. Diff.* 22 (1981), 45–51; MR 82g:13022.

[577] G. M. Kelly, A unified treatment of transfinite constructions for free algebras, free monoids, colimits, associated sheaves, and so on, *Bull. Austral. Math. Soc.* 22 (1980), 1–83; MR 82h:18003.

[578] G. M. Kelly, *Basic concepts of enriched category theory*, L.M.S. Lecture Notes Series no. 64 (Cambridge Univ. Press, 1982); MR 84e:18001.

[579] G. M. Kelly, Structures defined by finite limits in the enriched context, I, *Cahiers Top. Géom. Diff.* 23 (1982), 3–42; MR 83h:18007.

[580] G. M. Kelly, On the essentially-algebraic theory generated by a sketch, *Bull. Austral. Math. Soc.* 26 (1982), 45–56; MR 84g:18019.

[581] G. M. Kelly, Two addenda to the author's 'Transfinite Constructions', *Bull. Austral. Math. Soc.* 26 (1982), 221–237; MR 84a:18011.

[582] G. M. Kelly, A survey of totality for enriched and ordinary categories, *Cahiers Top. Géom. Diff. Catégoriques* 27 (1986), 109–132; MR 87k:18009.

[583] G. M. Kelly, Elementary observations on 2-categorical limits, *Bull. Austral. Math. Soc.* 39 (1989), 301–317; MR 90f:18004.

[584] G. M. Kelly, A note on relations relative to a factorization system, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 249–261; MR 93g:18007.

[585] G. M. Kelly, S. G. Lack and R. F. C. Walters, Coinverters and categories of fractions for categories with structure, *Appl. Categ. Struct.* 1 (1993), 95–102; MR 94k:18012.

[586] G. M. Kelly and F. W. Lawvere, On the complete lattice of essential localizations, *Bull. Soc. Math. Belg.* (A) 41 (1989), 289–319; MR 91c:18001.

[587] J. F. Kennison, On limit-preserving functors, *Illinois J. Math.* 12 (1968), 616–619; MR 38/208.

[588] J. F. Kennison, Integral domain type representations in sheaves and other topoi, *Math. Z.* 151 (1976), 35–56; MR 54/2758.

[589] J. F. Kennison, Galois theory and theaters of action in a topos, *J. Pure Appl. Alg.* 18 (1980), 149–164; MR 81j:18008.

[590] J. F. Kennison, Triples and compact sheaf representations, *J. Pure Appl. Alg.* 20 (1981), 13–38; MR 81m:18005.

[591] J. F. Kennison, Separable algebraic closure in a topos, *J. Pure Appl. Alg.* 24 (1982), 7–24; MR 84h:18007.

[592] J. F. Kennison, The fundamental group of a topos, *J. Pure Appl. Alg.* 30 (1983), 23–28; MR 85j:18007.

[593] J. F. Kennison, The fundamental group of a molecular topos, *J. Pure Appl. Alg.* 46 (1987), 187–215; MR 88k:18008.

[594] J. F. Kennison, What is the fundamental group?, *J. Pure Appl. Alg.* 59 (1989), 187–200; MR 90h:18003.

[595] J. F. Kennison, Pro-group actions and fundamental pro-groups, *J. Pure Appl. Alg.* 66 (1990), 185–218; MR 91j:18009.

[596] J. F. Kennison, The fundamental localic groupoid of a topos, *J. Pure Appl. Alg.* 77 (1992), 67–86; MR 93b:18009.

[597] J. F. Kennison, Synthetic solution manifolds for differential equations, *J. Pure Appl. Alg.* 143 (1999), 255–274; MR 00k:34022.

[598] J. F. Kennison and C. S. Ledbetter, Sheaf representations and the Dedekind reals, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer-Verlag, 1979), 500–513; MR 83j:18012.

[599] I. S. Kim, On some properties of the Blass topos, *J. Korea Soc. Math. Educ.* B 2 (1995), 25–29; MR 00c:18004.

[600] I. S. Kim, On the axiom of choice in a well-pointed topos, *J. Korea Soc. Math. Educ.* B 3 (1996), 131–139.

[601] I. S. Kim, A study of Boolean algebras in the Fraenkel–Mostowski topos, *J. Korea Soc. Math. Educ.* B7 (2000), 19–26; MR 01i:18008.

[602] B. Klunder, Topos based semantics for constructive logic with strong negation, *Pol. Acad. Sci. Inst. Philos. Sociol. Bull. Sect. Logic* 19 (1990), 133–138; MR 92b:03009; also in *Z. Math. Logik Grundl. Math.* 38 (1992), 509–519; MR 95c:03145.

[603] A. Kock, Bilinearity and cartesian closed monads, *Math. Scand.* 29 (1971), 161–174; MR 46/9141.

[604] A. Kock, Strong functors and monoidal monads, *Arch. Math. (Basel)* 23 (1972), 113–120; MR 46/3591.

[605] A. Kock, Monads for which structures are adjoint to units, preprint, Aarhus Univ. (1973); revised version published in *J. Pure Appl. Alg.* 104 (1995), 41–59; MR 96k:18009.

[606] A. Kock, Algebraic theories of power series, *Cahiers Top. Géom. Diff.* 16 (1975), 266–270.

[607] A. Kock, Linear algebra in a local ringed site, *Commun. Alg.* 3 (1975), 545–561; MR 52/5690.

[608] A. Kock, Universal projective geometry via topos theory, *J. Pure Appl. Alg.* 9 (1976), 1–24; MR 55/3959.

[609] A. Kock, A simple axiomatics for differentiation, *Math. Scand.* 40 (1977), 183–193; MR 56/12098.

[610] A. Kock, Taylor series calculus for ring objects of line type, *J. Pure Appl. Alg.* 12 (1978), 271–293; MR 58/18529.

[611] A. Kock, Formally real local rings, and infinitesimal stability, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 123–136; MR 81b:03064.

[612] A. Kock, On the synthetic theory of vector fields, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 139–157; MR 81c:58004.

[613] A. Kock, Formal manifolds, and synthetic theory of jet bundles, *Cahiers Top. Géom. Diff.* 21 (1980), 227–246; MR 82f:58006.

[614] A. Kock, Properties of well-adapted models for synthetic differential geometry, *J. Pure Appl. Alg.* 20 (1981), 55–70; MR 82f:58004.

[615] A. Kock, *Synthetic differential geometry*, L.M.S. Lecture Notes Series no. 51 (Cambridge Univ. Press, 1981); MR 83f:51023.

[616] A. Kock, Some problems and results in synthetic functional analysis, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 168–191; MR 85i:18004.

[617] A. Kock, Remarks on connections and sprays, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 192–202; MR 86k:53047.

[618] A. Kock, A combinatorial theory of connections, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 132–144; MR 85i:53077.

[619] A. Kock, Synthetic characterization of reduced algebras, *J. Pure Appl. Alg.* 36 (1985), 273–279; MR 87f:58003.

[620] A. Kock, Combinatorics of non-holonomous jets, *Czech. Math. J.* 35 (110) (1985), 419–428; MR 87c:58008.

[621] A. Kock, Combinatorial notions relating to principal fibre bundles, *J. Pure Appl. Alg.* 39 (1986), 141–151; MR 87c:18007.

[622] A. Kock, Convenient vector spaces embed into the Cahiers topos, *Cahiers Top. Géom. Diff. Catégoriques* 27 (1986), 3–17; MR 87m:58009; for corrigenda see [644].

[623] A. Kock, Synthetic reasoning in differential geometry, *Rev. Colomb. Mat.* 20 (1986), 179–185; MR 89h:58004.

[624] A. Kock, Lie-group-valued integration in well-adapted toposes, *Bull. Austral. Math. Soc.* 34 (1986), 395–410; MR 88d:58003.

[625] A. Kock, Generalized fibre bundles, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 194–207; MR 90c:18009.

[626] A. Kock, Fibre bundles in general categories, *J. Pure Appl. Alg.* 56 (1989), 233–245; MR 90b:18012.

[627] A. Kock, A Godement theorem for locales, *Math. Proc. Camb. Philos. Soc.* 105 (1989), 463–471; MR 90e:18005.

[628] A. Kock, On the integration theorem for Lie groupoids, *Czech. Math. J.* 39 (114) (1989), 423–431; MR 90j:20124.

[629] A. Kock, Algebras for the partial map classifier monad, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 262–278; MR 93g:18012.

[630] A. Kock, Combinatorics of curvature, and the Bianchi identity, *Theory Appl. Categ.* 2 (1996), 69–89; MR 98c:58001.

[631] A. Kock, Geometric construction of the Levi-Civita parallelism, *Theory Appl. Categ.* 4 (1998), 195–207; MR 00c:51017.

[632] A. Kock: The osculating plane of a space curve – synthetic formulations, *Rend. Circ. Mat. Palermo* (2) 64 (2000), 67–79; MR 01h:14035.

[633] A. Kock, Differential forms as infinitesimal cochains, *J. Pure Appl. Alg.* 154 (2000), 257–264; MR 01h:58002.

[634] A. Kock and R. Lavendhomme, Strong infinitesimal linearity, with applications to strong difference and affine connections, *Cahiers Top. Géom. Diff.* 25 (1984), 311–324; MR 87c:58004.

[635] A. Kock, P. Lecouturier and C. J. Mikkelsen, Some topos-theoretic concepts of finiteness, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 209–283; MR 52/2882.

[636] A. Kock and C. J. Mikkelsen, Topos-theoretic factorization of non-standard extensions, in *Proc. Victoria Symposium on Nonstandard Analysis*, Lecture Notes in Math. vol. 369 (Springer–Verlag, 1974), 122–143; MR 58/841.

[637] A. Kock and I. Moerdijk, Presentations of étendues, *Cahiers Top. Géom. Diff. Catégoriques* 32 (1991), 145–164; MR 92m:18007.

[638] A. Kock and I. Moerdijk, Every étendue comes from a local equivalence relation, *J. Pure Appl. Alg.* 82 (1992), 155–174; MR 93i:18005.

[639] A. Kock and T. Plewe, Glueing analysis for complemented subtoposes, *Theory Appl. Categ.* 2 (1996), 100–112; MR 97m:18008.

[640] A. Kock and G. E. Reyes, Doctrines in categorical logic, in *Handbook of mathematical logic*, Studies in Logic and the Foundations of Math. vol. 90 (North–Holland, 1977), 283–313; MR 58/10395.

[641] A. Kock and G. E. Reyes, Manifolds in formal differential geometry, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 514–533; MR 81d:58001.

[642] A. Kock and G. E. Reyes, Connections in formal differential geometry, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 158–195; MR 81h:53030.

[643] A. Kock and G. E. Reyes, Models for synthetic integration theory, *Math. Scand.* 48 (1981), 145–152; MR 83a:51023.

[644] A. Kock and G. E. Reyes, Corrigendum and addenda to the paper 'Convenient vector spaces embed into the Cahiers topos', *Cahiers Top. Géom. Diff. Catégoriques* 28 (1987), 99–110; MR 89d:58002.

[645] A. Kock and G. E. Reyes, Relatively Boolean and de Morgan toposes and locales, *Cahiers Top. Géom. Diff. Catégoriques* 35 (1994), 249–261; MR 95g:18004.

[646] A. Kock and G. E. Reyes, A note on frame distributions, *Cahiers Top. Géom. Diff. Catégoriques* 40 (1999), 127–140; MR 00i:18006.

[647] A. Kock and G. E. Reyes, Aspects of fractional exponent functors, *Theory Appl. Categ.* 5 (1999), 251–265; MR 00m:18009.

[648] A. Kock and J. Schmidt, A coherent theory of sites, *Bull. Soc. Math. Belg.* (A) 41 (1989), 321–331; MR 90m:18007.

[649] A. Kock and G. C. Wraith, *Elementary toposes*, Aarhus Univ. Lect. Notes Ser. 30 (1971); MR 49/7324.

[650] J. Koslowski, Continuous categories, in *Mathematical foundations of programming semantics*, Lecture Notes in Computer Science vol. 239 (Springer–Verlag, 1986), 149–161; MR 88b:18003.

[651] I. Kříž, A constructive proof of Tychonoff's theorem for locales, *Comment. Math. Univ. Carolinae* 26 (1985), 619–630; MR 87e:03113.

[652] I. Kříž and A. Pultr, Products of locally connected locales, *Rend. Circ. Mat. Palermo* (II) Suppl. 11 (1985), 61–70; MR 88j:54003.

[653] I. Kříž and A. Pultr, Peculiar behaviour of connected locales, *Cahiers Top. Géom. Diff. Catégoriques* 30 (1989), 25–43; MR 90e:54043.

[654] R. Kurata, Foundations of topoi, I: topoi in logic (Japanese), *Sûgaku* 35 (1983), 50–69; MR 84m:18006.

[655] C. Kuratowski, Sur la notion d'ensemble fini, *Fund. Math.* 1 (1920), 130–131.
[656] A. Labella, Costruzione del monoide dei quozienti in un topos elementare, *Rend. Mat.* (VI) 7 (1974), 151–168; MR 50/10009.
[657] G. Labelle, L'arithmétique toposienne des mots naturels, *Canad. Math. Bull.* 17 (1974), 685–688; MR 52/3103.
[658] C. Lair, Catégories modélables et catégories esquissables, *Diagrammes* 6 (1981), 1–120; MR 84j:18004.
[659] C. Lair, Sur le genre d'esquissabilité des catégories modelables (accessibles) possédant un objet terminal, *Diagrammes* 35 (1996), 3–23; MR 99a:18002a.
[660] C. Lair, Sur le genre d'esquissabilité des catégories modelables (accessibles) possédant les produits de deux, *Diagrammes* 35 (1996), 25–52; MR 99a:18002b.
[661] C. Lair, Sur le genre d'esquissabilité des catégories modelables (accessibles) possédant les limites d'indexations finies (resp. finies et non vides, finies et connexes, finies et connexes et non vides), *Diagrammes* 35 (1996), 53–90; MR 99a:18002c.
[662] F. Lamarche, On the stack completion of a category object in a pretopos, *C. R. Math. Rep. Acad. Sci. Canada* 10 (1988), 47–49; MR 89b:18007.
[663] J. Lambek, Deductive systems and categories, I, *Math. Systems Theory* 2 (1968), 287–318; MR 38/4277.
[664] J. Lambek, Deductive systems and categories, II, in *Category theory, homology theory and their applications*, vol. I, Lecture Notes in Math. vol. 86 (Springer-Verlag, 1969), 76–122; MR 39/3967.
[665] J. Lambek, Subequalizers, *Canad. Math. Bull.* 13 (1970), 337–349; MR 43/315.
[666] J. Lambek, Deductive systems and categories, III, in *Toposes, algebraic geometry and logic*, Lecture Notes in Math. vol. 274 (Springer-Verlag, 1972), 57–82; MR 50/1850.
[667] J. Lambek, From types to sets, *Adv. Math.* 36 (1980), 113–164; MR 81h:03124.
[668] J. Lambek, From λ-calculus to cartesian closed categories, in *To H. B. Curry: essays on combinatory logic, lambda calculus and formalism* (Academic Press, 1980), 375–402; MR 81m:03020.
[669] J. Lambek, Toposes are monadic over categories, in *Category theory*, Lecture Notes in Math. vol. 962 (Springer-Verlag, 1982), 153–166; MR 84b:18003.
[670] J. Lambek, Multicategories revisited, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 217–239; MR 90e:18010.
[671] J. Lambek, On the sheaf of possible worlds, in *Categorical topology* (World Scientific, 1989), 36–53; MR 91f:03128.
[672] J. Lambek, On some connections between logic and category theory, *Studia Logica* 48 (1989), 269–278; MR 91i:03126.
[673] J. Lambek, An equational variant of Lawvere's natural numbers object, *J. Pure Appl. Alg.* 154 (2000), 265–272; MR 01g:18010.
[674] J. Lambek and I. Moerdijk, Two sheaf representations of elementary toposes, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North-Holland, 1982), 275–295; MR 85h:03070.
[675] J. Lambek and B. A. Rattray, Localization and sheaf reflectors, *Trans. Amer. Math. Soc.* 210 (1975), 279–283; MR 56/5677.
[676] J. Lambek and P. J. Scott, Intuitionist type theory and the free topos, *J. Pure Appl. Alg.* 19 (1980), 215–257; MR 82b:03112.
[677] J. Lambek and P. J. Scott, Intuitionist type theory and foundations, *J. Philos. Logic* 10 (1981), 101–115; MR 83j:03111.
[678] J. Lambek and P. J. Scott, Algebraic aspects of topos theory, *Cahiers Top. Géom. Diff.* 22 (1981), 129–140; MR 84e:18005.

[679] J. Lambek and P. J. Scott, Independence of premisses and the free topos, in *Constructive mathematics*, Lecture Notes in Math. vol. 873 (Springer–Verlag, 1981), 191–207; MR 83e:03098.

[680] J. Lambek and P. J. Scott, New proofs of some intuitionistic principles, *Z. Math. Logik Grundl. Math.* 29 (1983), 493–504; MR 85e:03149.

[681] J. Lambek and P. J. Scott, Aspects of higher order categorical logic, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 145–174; MR 85i:03196.

[682] J. Lambek and P. J. Scott, *Introduction to higher order categorical logic*, Cambridge Studies in Advanced Math. no. 7 (Cambridge Univ. Press, 1986); MR 88a:03032.

[683] D. Lascar, On the category of models of a complete theory, *J. Symbolic Logic* 47 (1982), 249–266; MR 84g:03055.

[684] R. Lavendhomme, *Basic concepts of synthetic differential geometry* (Kluwer, 1996); MR 98e:58008.

[685] R. Lavendhomme and T. Lucas, Vers une géométrie algébrique locale, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 203–225; MR 85e:14004.

[686] R. Lavendhomme and T. Lucas, Une interprétation modale de la logique intuitionniste, *C. R. Acad. Sci. Paris* 298 (1984), I 193–196; MR 85b:03016.

[687] R. Lavendhomme and T. Lucas, A non-Boolean version of Feferman–Vaught's theorem, *Z. Math. Logik Grundl. Math.* 31 (1985), 299–308; MR 87d:03106.

[688] R. Lavendhomme and T. Lucas, Toposes and intuitionistic theories of types, *J. Pure Appl. Alg.* 57 (1989), 141–157; MR 90d:03138.

[689] R. Lavendhomme, T. Lucas and G. E. Reyes, Formal systems for topos-theoretic modalities, *Bull. Soc. Math. Belg.* (A) 41 (1989), 333–372; MR 91f:03032.

[690] R. Lavendhomme and H. Nishimura, Differential forms in synthetic differential geometry, *Internat. J. Theor. Phys.* 37 (1998), 2823–2832; MR 00h:58004.

[691] R. Lavendhomme and J. R. Roisin, Note on non-Abelian cohomology, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 534–541; MR 82k:18005.

[692] R. Lavendhomme and J. R. Roisin, Cohomologie non abélienne de structures algébriques, *J. Algebra* 67 (1980), 385–414; MR 82g:18004.

[693] J. D. Lawson, The duality of continuous posets, *Houston J. Math.* 5 (1979), 357–386; MR 81i:06003.

[694] F. W. Lawvere, Functorial semantics of algebraic theories, *Proc. Nat. Acad. Sci. U. S.A.* 50 (1963), 869–872; MR 28/2143.

[695] F. W. Lawvere, An elementary theory of the category of sets, *Proc. Nat. Acad. Sci. U. S.A.* 52 (1964), 1506–1511; MR 30/3025.

[696] F. W. Lawvere, Algebraic theories, algebraic categories and algebraic functors, in *The theory of models*, Studies in Logic and the Foundations of Math. (North-Holland, 1965), 413–418; MR 34/1373.

[697] F. W. Lawvere, The category of categories as a foundation for mathematics, in *Proc. Conf. on Categorical Algebra, La Jolla 1965* (Springer–Verlag, 1966), 1–20; MR 34/7332.

[698] F. W. Lawvere, Diagonal arguments and cartesian closed categories, in *Category theory, homology theory and their applications* vol. II, Lecture Notes in Math. vol. 92 (Springer–Verlag, 1969), 134–145; MR 39/4075.

[699] F. W. Lawvere, Adjointness in foundations, *Dialectica* 23 (1969), 281–296.

[700] F. W. Lawvere, Quantifiers and sheaves, in *Actes du Congrès International des Mathématiciens, Nice 1970*, Tome I (Gauthier–Villars, 1971), 329–334; MR 55/3029.

[701] F. W. Lawvere, Introduction, in *Toposes, algebraic geometry and logic*, Lecture Notes in Math. vol. 274 (Springer–Verlag, 1972), 1–12; MR 51/12793.

[702] F. W. Lawvere, Continuously variable sets: algebraic geometry = geometric logic, in *Logic Colloquium '73*, Studies in Logic and the Foundations of Math. vol. 80 (North-Holland 1975), 135–156; MR 52/5419.

[703] F. W. Lawvere, Introduction, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 3–14; MR 51/12982.

[704] F. W. Lawvere, Variable quantities and variable structures in topoi, in *Algebra, topology and category theory: a collection of papers in honor of Samuel Eilenberg* (Academic Press, 1976), 101–131; MR 54/7256.

[705] F. W. Lawvere, Categorical dynamics, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 1–28; MR 81c:18019.

[706] F. W. Lawvere, Toward the description in a smooth topos of the dynamically possible motions and deformations of a continuous body, *Cahiers Top. Géom. Diff.* 21 (1980), 377–392; MR 82i:58015.

[707] F. W. Lawvere, Taking categories seriously, *Rev. Colomb. Mat.* 20 (1986), 147–178; MR 89g:18016.

[708] F. W. Lawvere, Categories of spaces may not be generalized spaces, as exemplified by directed graphs, *Rev. Colomb. Mat.* 20 (1986), 179–186; MR 89h:18005.

[709] F. W. Lawvere, Qualitative distinctions between some toposes of generalized graphs, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 261–299; MR 90g:18005.

[710] F. W. Lawvere, Display of graphics and their applications, as exemplified by 2-categories and the Hegelian 'taco', in *Proc. First Internat. Conf. on Algebraic Methodology and Software Technology* (Univ. of Iowa, 1989), 51–75.

[711] F. W. Lawvere, More on graphic toposes, *Cahiers Top. Géom. Diff. Catégoriques* 32 (1991), 5–10; MR 92m:18008.

[712] F. W. Lawvere, Some thoughts on the future of category theory, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 1–13; MR 93g:18001.

[713] F. W. Lawvere, Intrinsic co-Heyting boundaries and the Leibniz rule in certain toposes, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 279–281; MR 93e:18007.

[714] F. W. Lawvere, Cohesive toposes and Cantor's 'lauter Einsen', *Philos. Math.* (3) 2 (1994), 5–15; MR 95e:00020.

[715] F. W. Lawvere, Unity and identity of opposites in calculus and physics, *Appl. Categ. Struct.* 4 (1996), 167–174; MR 97g:18003.

[716] F. W. Lawvere, Comments on the development of topos theory, in *The development of mathematics 1950–2000* (Birkhäuser–Verlag, 2000), 715–734; MR 01i:18009.

[717] F. W. Lawvere and S. H. Schanuel, *Conceptual mathematics* (Cambridge Univ. Press, 1997); MR 99e:18001.

[718] O. Leroy, Groupoïde fondamental et théorème de van Kampen en théorie des topos, *Cahiers Math. Montpellier* 17 (1979); MR 82i:18006.

[719] B. Lesaffre, Structures algébriques dans les topos élémentaires, *C. R. Acad. Sci. Paris* 277 (1973), A663–666; MR 48/8597.

[720] B. Lesaffre, Théories algébriques internes dans les topos, *C. R. Acad. Sci. Paris* 286 (1978), A1103–1106; MR 58/11062.

[721] D. B. Lever, Continuous families: categorical aspects, *Cahiers Top. Géom. Diff.* 24 (1983), 393–432; MR 86b:18004.

[722] D. B. Lever, Precategory objects of toposes, *J. Pure Appl. Alg.* 35 (1985), 47–75; MR 86h:18001.

[723] L. G. Lewis, Open maps, colimits, and a convenient category of fibre spaces, *Topology Appl.* 19 (1985), 75–89; MR 86j:54049.

[724] F. E. J. Linton, Some aspects of equational categories, in *Proceedings of the Conference on Categorical Algebra, La Jolla 1965* (Springer–Verlag, 1966), 84–94; MR 35/233.

[725] F. E. J. Linton, Autonomous equational categories, *J. Math. Mech.* 15 (1966), 637–642; MR 32/7619.

[726] F. E. J. Linton, An outline of functorial semantics, in *Seminar on triples and categorical homology theory*, Lecture Notes in Math. vol. 80 (Springer–Verlag, 1969), 7–52; MR 39/5655.

[727] F. E. J. Linton, Coequalizers in categories of algebras, in *Seminar on triples and categorical homology theory*, Lecture Notes in Math. vol. 80 (Springer–Verlag, 1969), 75–90; MR 39/5656.

[728] F. E. J. Linton, Injectives in topoi, III: stability under coproducts, *Bull. Acad. Polon. Sci.* 29 (1981), 341–347; MR 83j:18007.

[729] F. E. J. Linton and R. Paré, Injectives in topoi, I: representing coalgebras as algebras, in *Categorical topology*, Lecture Notes in Math. vol. 719 (Springer–Verlag, 1979), 196–206; MR 80j:18007a.

[730] B. Loiseau, Some higher-order properties of a natural number object in a topos, *Bull. Soc. Math. Belg.* (B) 42 (1990), 361–376; MR 91i:03127.

[731] B. Loiseau and M. M. Mawanda, On natural number objects, finiteness and Kripke–Platek models in toposes, *J. Pure Appl. Alg.* 61 (1989), 257–266; MR 90k:18005.

[732] J. R. Longley and A. K. Simpson, A uniform approach to domain theory in realizability models, *Math. Struct. Comput. Sci.* 7 (1997), 469–505; MR 98m:68168.

[733] G. Loullis, Infinite forcing for Boolean valued models, *Bull. Greek Math. Soc.* 19 (1978), 155–182; MR 80f:03042.

[734] G. Loullis, Sheaves and Boolean valued model theory, *J. Symbolic Logic* 44 (1979), 153–183; MR 80k:03037.

[735] V. A. Lyubetskiĭ, Estimates and sheaves: some problems in nonstandard analysis (Russian), *Uspekhi Mat. Nauk* 44/4 (1989), 99–153; English translation in *Russian Math. Surveys* 44/4 (1989), 37–112; MR 91k:03155.

[736] V. A. Lyubetskiĭ and E. I. Gordon, Embedding of sheaves in a Heyting-valued universe (Russian), *Dokl. Akad. Nauk. SSSR* 268 (1983), 794–798; English translation in *Soviet Math. Dokl.* 27 (1982), 159–164; MR 84f:03034.

[737] W. MacCaull, On the validity of Hilbert's Nullstellensatz, Artin's theorem and related results in Grothendieck toposes, *J. Symbolic Logic* 53 (1988), 1177–1187; MR 90a:03109.

[738] W. MacCaull, Hilbert's Nullstellensatz revisited, *J. Pure Appl. Alg.* 54 (1988), 289–297; MR 90c:18003.

[739] W. MacCaull, Positive definite functions over regular f-rings and representations as sums of squares, *Ann. Pure Appl. Logic* 44 (1989), 243–257; MR 91a:03085.

[740] J. MacDonald and A. Stone, Topoi over graphs, *Cahiers Top. Géom. Diff.* 25 (1984), 51–62; MR 86g:18002.

[741] A. J. Macintyre, Model completeness for sheaves of structures, *Fund. Math.* 81 (1973), 73–89; MR 49/2361.

[742] S. Mac Lane, *Categories for the working mathematician*, Graduate Texts in Math. no. 5 (Springer–Verlag, 1971); MR 50/7275; revised edition 1998; MR 01j:18001.

[743] S. Mac Lane, Sets, topoi and internal logic in categories, in *Logic Colloquium '73*, Studies in Logic and the Foundations of Math. vol. 80 (North-Holland, 1975), 119–134; MR 52/5419.
[744] S. Mac Lane, The genesis of mathematical structures, as exemplified in the work of Charles Ehresmann, *Cahiers Top. Géom. Diff.* 21 (1980), 353–365; MR 83b:01056.
[745] S. Mac Lane, History of abstract algebra: origins, rise and decline of a movement, in *American mathematical heritage: algebra and applied mathematics*, Math. Ser. no. 13 (Texas Tech. Univ., 1981), 3–35; MR 83m:01042.
[746] S. Mac Lane, The development of mathematical ideas by collision: the case of categories and topos theory, in *Categorical topology* (World Scientific, 1989), 1–9; MR 91e:18001.
[747] S. Mac Lane, Categories in geometry, algebra and logic, *Math. Japon.* 42 (1995), 169–178; MR 96e:18001.
[748] S. Mac Lane, The development and prospects for category theory, *Appl. Categ. Struct.* 4 (1996), 129–136; MR 97e:18001.
[749] S. Mac Lane, Samuel Eilenberg's work in category theory, *Appl. Categ. Struct.* 4 (1996), 137–138.
[750] S. Mac Lane and I. Moerdijk, Reals and forcing with an elementary topos, in *Logic from computer science*, MSRI Publ. no. 21 (Springer-Verlag, 1992), 373–385; MR 94h:03133.
[751] S. Mac Lane and I. Moerdijk, *Sheaves in geometry and logic: a first introduction to topos tTheory* (Springer-Verlag, 1992); MR 96c:03119.
[752] S. Mac Lane and I. Moerdijk, Topos theory, in *Handbook of algebra*, vol. I (North-Holland, 1996), 501–528; MR 97j:18005.
[753] C. McLarty, Local, and some global, results in synthetic differential geometry, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 226–256.
[754] C. McLarty, Left exact logic, *J. Pure Appl. Alg.* 41 (1986), 63–66; MR 88a:03163.
[755] C. McLarty, Elementary axioms for canonical points of toposes, *J. Symbolic Logic* 52 (1987), 202–204; MR 88a:03164.
[756] C. McLarty, Defining sets as sets of points of spaces, *J. Philos. Logic* 17 (1988), 75–90; MR 88m:03095.
[757] C. McLarty, Stable surjection logic, *Diagrammes* 22 (1989), 45–57; MR 92a:03079.
[758] C. McLarty, The uses and abuses of the history of topos theory, *Brit. J. Philos. Sci.* 41 (1990), 351–375; MR 92b:01044.
[759] C. McLarty, Axiomatizing a category of categories, *J. Symbolic Logic* 56 (1991), 1243–1260; MR 93a:03010.
[760] C. McLarty, *Elementary categories, elementary toposes*, Oxford Logic Guides no. 21 (Oxford Univ. Press, 1992); MR 94c:18001.
[761] C. McLarty, Failure of cartesian closedness in NF, *J. Symbolic Logic* 57 (1992), 555–556; MR 94a:03084.
[762] C. McLarty, Category theory in real time, *Philos. Math.* (3) 2 (1994), 36–44; MR 95d:00005.
[763] R. G. McLean and M. W. Warner, Locale theory and fuzzy topology, *Fuzzy Sets Syst.* 54 (1993), 91–97; MR 94a:54027.
[764] D. S. Macnab, Some applications of double-negation sheafification, *Proc. Edinburgh Math. Soc.* 20 (1977), 279–285; MR 57/412.
[765] D. S. Macnab, Modal operators on Heyting algebras, *Alg. Universalis* 12 (1981), 5–29; MR 82j:03082.
[766] J. J. Madden, κ-frames, *J. Pure Appl. Alg.* 70 (1991), 107–127; MR 92d:06030.

[767] J. J. Madden and A. Molitor, Epimorphisms of frames, *J. Pure Appl. Alg.* 70 (1991), 129–132; MR 92d:06031.
[768] J. J. Madden and J. Vermeer, Lindelöf locales and realcompactness, *Math. Proc. Camb. Philos. Soc.* 99 (1986), 473–480; MR 87g:54060.
[769] L. Mahé, Topos infinitesimal, *C. R. Acad. Sci. Paris* 277 (1973), A497–500; MR 49/8986.
[770] L. Mahé, Théorèmes, exemples et contre-exemples pour certaines propriétés des entiers naturels dans les topos, *C. R. Acad. Sci. Paris* 282 (1976), A1273–1276; MR 54/7255.
[771] M. E. Maietti, Some aspects of the categorical semantics for the polymorphic λ-calculus, in *Logic and algebra*, Lecture Notes Pure Appl. Math. vol. 180 (Marcel Dekker, 1996), 589–601; MR 99a:03022.
[772] M. E. Maietti, The internal type theory of a Heyting pretopos, in *Types for proofs and programs*, Lecture Notes in Computer Science vol. 1512 (Springer–Verlag, 1998), 216–235.
[773] M. E. Maietti, V. C.V. de Paiva and E. Ritter, Categorical models for intuitionistic and linear type theory, in *Foundations of software science and computer structures*, Lecture Notes in Computer Science vol. 1784 (Springer–Verlag, 2000), 223–237; MR 01i:03131.
[774] M. E. Maietti and S. Valentini, Can you add power-sets to Martin-Löf's intuitionistic set theory?, *Math. Logic Quarterly* 45 (1999), 521–532; MR 00h:03020.
[775] M. Makkai, On full embeddings, I, *J. Pure Appl. Alg.* 16 (1980), 183–195; MR 83g:18007.
[776] M. Makkai, The topos of types, in *Logic year 1979–80*, Lecture Notes in Math. vol. 859 (Springer–Verlag, 1981), 157–201; MR 82k:03103.
[777] M. Makkai, Full continuous embeddings of toposes, *Trans. Amer. Math. Soc.* 269 (1982), 167–196; MR 83c:03058.
[778] M. Makkai, A Stone-type representation theory for first-order logic, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 175–243; MR 85i:03197.
[779] M. Makkai, Ultraproducts and categorical logic, in *Methods in mathematical logic*, Lecture Notes in Math. vol. 1130 (Springer–Verlag, 1985), 222–309; MR 87f:03174.
[780] M. Makkai, Stone duality for first order logic, *Adv. Math.* 65 (1987), 97–170; MR 89h:03067.
[781] M. Makkai, Strong conceptual completeness for first order logic, *Ann. Pure Appl. Logic* 40 (1988), 167–215; MR 90b:03096.
[782] M. Makkai, A theorem on Barr-exact categories, with an infinitary generalization, *Ann. Pure Appl. Logic* 47 (1990), 225–268; MR 91h:18003.
[783] M. Makkai, *Duality and definability in first order logic*, Memoirs Amer. Math. Soc. no. 503 (1993); MR 94b:03104.
[784] M. Makkai, Avoiding the axiom of choice in general category theory, *J. Pure Appl. Alg.* 108 (1996), 109–173; MR 97h:18002.
[785] M. Makkai, Generalized sketches as a framework for completeness theorems, *J. Pure Appl. Alg.* 115 (1997), 49–79, 179–212 and 241–274; MR 98a:03103–5.
[786] M. Makkai, Towards a categorical foundation of mathematics, in *Logic Colloquium '95*, Lecture Notes in Logic vol. 11 (Springer–Verlag, 1998), 153–190; MR 00f:03015.
[787] M. Makkai and R. Paré, *Accessible categories: the foundations of categorical model theory*, Contemp. Math. vol. 104 (Amer. Math. Soc., 1989); MR 91a:03118.

[788] M. Makkai and A. M. Pitts, Some results on locally finitely presentable categories, *Trans. Amer. Math. Soc.* 299 (1987), 473–496; MR 88a:03162.

[789] M. Makkai and G. E. Reyes, Model-theoretic methods in the theory of topoi and related categories, *Bull. Acad. Polon. Sci.* 24 (1976), 379–392; MR 54/10009.

[790] M. Makkai and G. E. Reyes, *First order categorical logic*, Lecture Notes in Math. vol. 611 (Springer–Verlag, 1977); MR 58/21600.

[791] M. Makkai and G. E. Reyes, Completeness results for intuitionistic and modal logic in a categorical setting, *Ann. Pure Appl. Logic* 72 (1995), 25–101; MR 96b:03087.

[792] J. Malgoire and C. Voisin, Factorisation de Stein topologique (découpe), *Topology* 20 (1981), 191–207; MR 83m:54015.

[793] J. Malitz, The Hanf number for complete $L_{\omega_1\omega}$ sentences, in *The syntax and semantics of infinitary languages*, Lecture Notes in Math. vol. 72 (Springer–Verlag, 1968), 166–181.

[794] C. Marchini, Funtori che conservano e riflettono le formule di Horn, *Atti Sem. Mat. Fis. Univ. Modena* 22 (1973), 60–65; MR 51/3251.

[795] C. Marchini, Funtori rappresentabili e formule di Horn, *Atti Sem. Mat. Fis. Univ. Modena* 22 (1973), 335–347; MR 52/7891.

[796] C. Marchini, Alcune questioni di semantica in un topos elementare, *Riv. Mat. Univ. Parma* (4) 2 (1976), 213–220; MR 56/80.

[797] C. Marchini, Confronto tra due semantiche categoriali, *Boll. Un. Mat. Ital.* A (5) 16 (1979), 103–108; MR 80m:03113.

[798] F. Marmolejo, Continuous families of coalgebras, *J. Pure Appl. Alg.* 130 (1998), 197–215; MR 99f:18015.

[799] P. Martin-Löf, An intuitionistic theory of types: predicative part, in *Logic Colloquium '73*, Studies in Logic and the Foundations of Math. vol. 80 (North-Holland, 1975), 73–118; MR 52/7856; revised version in *Twenty-five years of constructive type theory*, Oxford Logic Guides no. 36 (Oxford Univ. Press, 1998), 127–172.

[800] P. Martin-Löf, *Intuitionistic type theory* (Bibliopolis, 1984); MR 86j:03005.

[801] C. Maurer, Universes in topoi, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 284–296; MR 52/95.

[802] M. M. Mawanda, Well-ordering and choice in toposes, *J. Pure Appl. Alg.* 50 (1988), 171–184; MR 89f:03062.

[803] M. M. Mawanda, Finite cardinals in general topoi, *J. Pure Appl. Alg.* 50 (1988), 185–194; MR 89f:03063.

[804] X. Q. Meng, A brief introduction to topos theory (Chinese), *Adv. Math. (Beijing)* 21 (1992), 1–24.

[805] B. N. Mesablishvili, Fundamental theorem for finite Galois extensions of an internal connected commutative ring in an elementary topos, and the functor T (Russian), *Tbiliss. Gos. Univ. Inst. Prikl. Mat. Trudy* 36 (1990), 9–27; MR 92j:18006.

[806] E. Michael, Bi-quotient maps and cartesian products of quotient maps, *Ann. Inst. Fourier* 18 (1968), 287–302; MR 39/6277.

[807] E. Michael, Complete spaces and tri-quotient maps, *Illinois J. Math.* 21 (1977), 716–733; MR 57/7543.

[808] M. V. Mielke, Convenient categories for internal singular algebraic topology, *Illinois J. Math.* 27 (1983), 519–534; MR 84m:54011.

[809] M. V. Mielke, Homotopically trivial toposes, *Pacific J. Math.* 110 (1984), 171–182; MR 85b:18004.

[810] M. V. Mielke, Topos based homology theory, *Comment. Math. Univ. Carolinae* 34 (1993), 549–565; MR 94e:55010.

[811] C. J. Mikkelsen, *Lattice-theoretic and logical aspects of elementary topoi*, Aarhus Universitet Var. Publ. Ser. 25 (1976); MR 55/2572.

[812] K. C. Min, Y. S. Kim and J. W. Park, Fibrewise exponential laws in a quasitopos, *Cahiers Top. Géom. Diff.* 40 (1999), 242–260; MR 01c:18004.

[813] M. C. Minguez, Wedge product of forms in synthetic differential geometry, *Cahiers Top. Géom. Diff. Catégoriques* 29 (1988), 59–66; MR 89h:58006.

[814] F. Miraglia, The downward Löwenheim–Skolem theorem for L-structures in Ω-sets, in *Methods and applications of mathematical logic*, Contemp. Math. vol. 69 (Amer. Math. Soc., 1988), 189–208; MR 89h:03059.

[815] W. Mitchell, Boolean topoi and the theory of sets, *J. Pure Appl. Alg.* 2 (1972), 261–274; MR 47/8299.

[816] W. Mitchell, On topoi as closed categories, *J. Pure Appl. Alg.* 3 (1973), 133–139; MR 50/427.

[817] W. Mitchell, Categories of Boolean topoi, *J. Pure Appl. Alg.* 3 (1973), 193–201; MR 50/9585.

[818] A. Möbus, Syntactical theory of functors, *Cahiers Top. Géom. Diff.* 23 (1982), 47–53; MR 84a:03081.

[819] J. Močkoř, Sheaves and approximation theorems, *Comment. Math. Univ. St. Paul.* 38 (1989), 199–212; MR 90k:18013.

[820] J. L. Moens, *Caractérisation des topos de faisceaux sur un site interne à un topos*, Ph.D. thesis, Univ. Catholique de Louvain (1982).

[821] I. Moerdijk, Glueing topoi and higher order disjunction and existence, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North-Holland, 1982), 359–375; MR 84k:03149.

[822] I. Moerdijk, Some topological spaces which are universal for intuitionistic predicate logic, *Proc. Kon. Ned. Akad. Weten.* (A) 85 = *Indag. Math.* 44 (1982), 227–235; MR 83j:03023.

[823] I. Moerdijk, Equational classes of toposes, *Commun. Alg.* 11 (1983), 839–842; MR 84g:03106.

[824] I. Moerdijk, On the Freyd cover of a topos, *Notre Dame J. Formal Logic* 24 (1983), 517–526; MR 85h:03072.

[825] I. Moerdijk, Spaced spaces, *Compositio Math.* 53 (1984), 171–209; MR 87a:18007.

[826] I. Moerdijk, Heine–Borel does not imply the Fan theorem, *J. Symbolic Logic* 49 (1984), 514–519; MR 86h:03106.

[827] I. Moerdijk, An elementary proof of the descent theorem for Grothendieck toposes, *J. Pure Appl. Alg.* 37 (1985), 185–191; MR 86k:18002.

[828] I. Moerdijk, Continuous fibrations and inverse limits of toposes, *Compositio Math.* 58 (1986), 45–72; MR 87f:18004.

[829] I. Moerdijk, On the embedding of manifolds into the smooth Zariski topos, in *Mathematical logic and theoretical computer science*, Lecture Notes Pure Appl. Math. vol. 106 (Marcel Dekker, 1987), 261–276; MR 89c:18008.

[830] I. Moerdijk, Path-lifting for Grothendieck toposes, *Proc. Amer. Math. Soc.* 102 (1988), 242–248; MR 89c:18004; addendum *ibid.* 125 (1997), 2815–2818; MR 97j:18006.

[831] I. Moerdijk, The classifying topos of a continuous groupoid, I, *Trans. Amer. Math. Soc.* 310 (1988), 629–668; MR 90a:18005.

[832] I. Moerdijk, Toposes and groupoids, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer-Verlag, 1988), 280–298; MR 89m:18003.

[833] I. Moerdijk, Morita equivalence for continuous groups, *Math. Proc. Camb. Philos. Soc.* 103 (1988), 97–115; MR 88j:22001.

[834] I. Moerdijk, Prodiscrete groups and Galois toposes, *Proc. Kon. Nederl. Akad. Weten.* (A) 92 = *Indag. Math.* 51 (1989), 219–234; MR 90h:18014.
[835] I. Moerdijk, Descent theory for toposes, *Bull. Soc. Math. Belg.* (A) 41 (1989), 373–391; MR 91b:18005.
[836] I. Moerdijk, Bisimplicial sets and the group-completion theorem, in *Algebraic K-theory: connections with geometry and topology*, NATO Adv. Sci. Inst. Ser. C 279 (Kluwer, 1989), 225–240; MR 91j:18021.
[837] I. Moerdijk, The classifying topos of a continuous groupoid, II, *Cahiers Top. Géom. Diff. Catégoriques* 31 (1990), 137–168; MR 92c:18003.
[838] I. Moerdijk, Classifying toposes and foliations, *Ann. Inst. Fourier* 41 (1991), 189–209; MR 92i:57028.
[839] I. Moerdijk, Foliations, groupoids and Grothendieck étendues, *Rev. Acad. Cienc. Zaragoza* (2) 48 (1993), 5–33; MR 95d:58148.
[840] I. Moerdijk, Espace classifiant, topos classifiant et ordres augmentées, *C. R. Acad. Sci. Paris* 317 (1993), I 293–298; MR 94d:18005.
[841] I. Moerdijk, A model for intuitionistic non-standard arithmetic, *Ann. Pure Appl. Logic* 73 (1995), 37–51; MR 96k:03145.
[842] I. Moerdijk, *Classifying spaces and classifying topoi*, Lecture Notes in Math. vol. 1616 (Springer–Verlag, 1995); MR 98f:18006.
[843] I. Moerdijk, Classifying spaces for toposes with enough points, *Rend. Sem. Mat. Fis. Milano* 66 (1996), 377–389; MR 99m:18002.
[844] I. Moerdijk, Proof of a conjecture of A. Haefliger, *Topology* 37 (1998), 735–741; MR 99b:18015.
[845] I. Moerdijk, Sets, topoi, and intuitionism, *Philos. Math.* (3) 6 (1998), 169–177; MR 99k:03009.
[846] I. Moerdijk and E. Palmgren, Minimal models of Heyting arithmetic, *J. Symbolic Logic* 62 (1997), 1448–1460; MR 99i:03074.
[847] I. Moerdijk and E. Palmgren, Wellfounded trees in categories, *Ann. Pure Appl. Logic* 104 (2000), 189–218; MR 01j:03016.
[848] I. Moerdijk and D. A. Pronk, Orbifolds, sheaves and groupoids, *K-Theory* 12 (1997), 13–21; MR 98i:22004.
[849] I. Moerdijk, N. V. Quê and G. E. Reyes, Forcing smooth square roots and integration, *Fund. Math.* 128 (1987), 145–158; MR 89d:58003.
[850] I. Moerdijk and G. E. Reyes, Cohomology theories in synthetic differential geometry, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 1–67; MR 85m:58004; also in *Rev. Colomb. Mat.* 20 (1986), 223–276; MR 89k:58004.
[851] I. Moerdijk and G. E. Reyes, Smooth spaces versus continuous spaces in models for synthetic differential geometry, *J. Pure Appl. Alg.* 32 (1984), 143–176; MR 85i:03198.
[852] I. Moerdijk and G. E. Reyes, De Rham's theorem in a smooth topos, *Math. Proc. Camb. Philos. Soc.* 96 (1984), 61–71; MR 85e:18001.
[853] I. Moerdijk and G. E. Reyes, Rings of smooth functions and their localizations, I, *J. Algebra* 99 (1986), 324–336; MR 87j:58010.
[854] I. Moerdijk and G. E. Reyes, On the relation between connections and sprays, *Rev. Colomb. Mat.* 20 (1986), 187–222; MR 89k:18019.
[855] I. Moerdijk and G. E. Reyes, A smooth version of the Zariski topos. *Adv. Math.* 65 (1987), 229–253; MR 88k:18013.
[856] I. Moerdijk and G. E. Reyes, *Models for smooth infinitesimal analysis* (Springer–Verlag, 1991); MR 91m:18017.
[857] I. Moerdijk and J. J. C. Vermeulen, Proof of a conjecture of A. Pitts, *J. Pure Appl. Alg.* 143 (1999), 329–338; MR 01c:18003.

[858] I. Moerdijk and J. J. C. Vermeulen, *Proper maps of toposes*, Memoirs Amer. Math. Soc. no. 705 (2000); MR 01j:18008.
[859] I. Moerdijk and G. C. Wraith, Connected locally connected toposes are path-connected, *Trans. Amer. Math. Soc.* 295 (1986), 849–859; MR 87c:18002.
[860] G. P. Monro, On generic extensions without the axiom of choice, *J. Symbolic Logic* 48 (1983), 39–52; MR 84d:03061.
[861] G. P. Monro, Quasitopoi, logic and Heyting-valued models, *J. Pure Appl. Alg.* 42 (1986), 141–164; MR 88m:03096.
[862] G. P. Monro, A category-theoretic approach to Boolean-valued models of set theory, *J. Pure Appl. Alg.* 42 (1986), 245–274; MR 88f:03045.
[863] G. P. Monro, Logic, sheaves and factorization systems, *J. Symbolic Logic* 58 (1993), 872–893; MR 94k:03082.
[864] P. S. Mulry, Generalized Banach–Mazur functionals in the topos of recursive sets, *J. Pure Appl. Alg.* 26 (1982), 71–83; MR 84e:03079.
[865] P. S. Mulry, Some connections between models of computation, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 317–324; MR 91a:68061.
[866] P. S. Mulry, A categorical approach to the theory of computation, *Ann. Pure Appl. Logic* 43 (1989), 293–305; MR 91f:03129.
[867] P. S. Mulry, Categorical fixed point semantics, *Theoret. Comput. Sci.* 70 (1990), 85–97; MR 91c:03049.
[868] P. S. Mulry, Partial map classifiers and partial cartesian closed categories, *Theoret. Comput. Sci.* 136 (1994), 109–123; MR 95k:68128.
[869] C. J. Mulvey, Intuitionistic algebra and representations of rings, in *Recent advances in the representation theory of rings and C^*-algebras by continuous sections*, Memoirs Amer. Math. Soc. no. 148 (1974), 3–57; MR 53/2650.
[870] C. J. Mulvey, A generalization of Swan's theorem, *Math. Z.* 151 (1976), 51–70; MR 55/2909.
[871] C. J. Mulvey, A remark on the prime stalk theorem, *J. Pure Appl. Alg.* 10 (1977), 253–256; MR 57/397.
[872] C. J. Mulvey, A categorical characterization of compactness, *J. London Math. Soc.* (2) 17 (1978), 356–362; MR 57/10673.
[873] C. J. Mulvey, Compact ringed spaces, *J. Algebra* 52 (1978), 411–436; MR 80b:54016.
[874] C. J. Mulvey, A generalization of Gelfand duality, *J. Algebra* 56 (1979), 499–505; MR 80d:16028.
[875] C. J. Mulvey, Banach spaces over a compact space, in *Categorical topology*, Lecture Notes in Math. vol. 719 (Springer–Verlag, 1979), 243–249; MR 80m:18005.
[876] C. J. Mulvey, Representations of rings and modules, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 542–585; MR 82e:16039.
[877] C. J. Mulvey, Banach sheaves, *J. Pure Appl. Alg.* 17 (1980), 69–83; MR 81h:46097.
[878] C. J. Mulvey, A non-commutative Gel'fand–Naimark theorem, *Math. Proc. Camb. Philos. Soc.* 88 (1980), 425–428; MR 82m:46062.
[879] C. J. Mulvey, &, *Rend. Circ. Mat. Palermo Suppl.* (II) 12 (1986), 99–104; MR 87j:81017.
[880] C. J. Mulvey, The maximality of filters, *J. Pure Appl. Alg.* 68 (1990), 253–258; MR 91m:06022.
[881] C. J. Mulvey and J. W. Pelletier, The dual locale of a seminormed space, *Cahiers Top. Géom. Diff.* 23 (1982), 73–92; MR 84k:46052.

[882] C. J. Mulvey and J. W. Pelletier, On points of locales in a De Morgan topos, in *Categorical topology*, Sigma Ser. Pure Math. 5 (Heldermann–Verlag, 1984), 392–407; MR 86j:18005.

[883] C. J. Mulvey and J. W. Pelletier, A globalisation of the Hahn–Banach theorem, *Adv. Math.* 89 (1991), 1–59; MR 93c:46132.

[884] C. J. Mulvey and J. W. Pelletier, A quantisation of the calculus of relations, in *Category theory 1991*, C.M.S. Conf. Proc. vol. 13 (Amer. Math. Soc., 1992), 345–360; MR 94c:18016.

[885] G. S. Murchiston, Note on paper of Geanău, *Bull. Math. Soc. Sci. Math. Roum.* 27 (75) (1983), 73–74; MR 84d:18004.

[886] G. S. Murchiston and M. G. Stanley, A "T_1" space with no closed points, and a "T_1" locale which is not 'T_1', *Math. Proc. Camb. Philos. Soc.* 95 (1984), 421–422; MR 86d:54004.

[887] C. V. Negoiţă, Fuzzy sets in topoi, *Fuzzy Sets Syst.* 8 (1982), 93–99; MR 84e:03080.

[888] C. V. Negoiţă and A. C. Ştefănescu, Fuzzy objects in topoi: a generalization of fuzzy sets, *Bul. Inst. Politehn. Iaşi* (I) 24 (1978), 25–28; MR 80j:03081.

[889] S. Negri, Stone bases, alias the constructive content of Stone representation, in *Logic and algebra*, Lecture Notes Pure Appl. Math. vol. 180 (Marcel Dekker, 1996), 617–636; MR 97f:06023.

[890] S. Negri and D. Soravia, The continuum as a formal space, *Arch. Math. Logic* 38 (1999), 423–447; MR 00k:03144.

[891] S. Negri and S. Valentini, Tychonoff's theorem in the framework of formal topologies, *J. Symbolic Logic* 62 (1997), 1315–1332; MR 99d:54005.

[892] S. B. Niefield, Cartesian inclusions: locales and toposes, *Commun. Alg.* 16 (1981), 1639–1671; MR 83a:18012.

[893] S. B. Niefield, Cartesianness: topological spaces, uniform spaces and affine schemes, *J. Pure Appl. Alg.* 23 (1982), 147–167; MR 83b:18012.

[894] S. B. Niefield, Cartesian spaces over T and locales over $\Omega(T)$, *Cahiers Top. Géom. Diff.* 23 (1982), 257–267; MR 84c:54016.

[895] S. B. Niefield, Exactness and projectivity, in *Category theory*, Lecture Notes in Math. vol. 962 (Springer–Verlag, 1982), 221–227; MR 84c:18016.

[896] S. B. Niefield, Exponentiable morphisms: posets, spaces, locales and Grothendieck toposes, *Theory Appl. Categ.* 8 (2001), 16–32; MR 01k:18007.

[897] S. B. Niefield and K. I. Rosenthal, Spec(R), ideals and the strong De Morgan's law, *J. Algebra* 93 (1985), 169–181; MR 86f:18009.

[898] S. B. Niefield and K. I. Rosenthal, A note on the algebraic De Morgan's law, *Cahiers Top. Géom. Diff.* 26 (1985), 115–120; MR 86m:13003.

[899] S. B. Niefield and K. I. Rosenthal, Sheaves of integral domains on Stone spaces, *J. Pure Appl. Alg.* 47 (1987), 173–179; MR 88m:18015.

[900] S. B. Niefield and K. I. Rosenthal, Spatial sublocales and essential primes, *Topology Appl.* 26 (1987), 263–269; MR 88k:54023.

[901] S. B. Niefield and K. I. Rosenthal, Constructing locales from quantales, *Math. Proc. Camb. Philos. Soc.* 104 (1988), 215–234; MR 89f:18003.

[902] S. B. Niefield and K. I. Rosenthal, Componental nuclei, in *Categorical algebra and its applications*, Lecture Notes in Math. vol. 1348 (Springer–Verlag, 1988), 299–306; MR 90a:06013.

[903] S. B. Niefield and K. I. Rosenthal, Sheaves of semiprime ideals, *Cahiers Top. Géom. Diff. Catégoriques* 31 (1990), 213–228; MR 92b:18003.

[904] S. B. Niefield and S. H. Sun, Algebraic De Morgan's laws for non-commutative rings, *Cahiers Top. Géom. Diff. Catégoriques* 36 (1995), 127–140; MR 96c:16006.

[905] H. Nishimura, Logical quantizations of first-order structures, *Internat. J. Theor. Phys.* 35 (1996), 495–517; MR 98g:18005.

[906] H. Nishimura, Logical quantization of topos theory, *Internat. J. Theor. Phys* 35 (1996), 2555–2596; MR 98j:81028.

[907] H. Nishimura, Nonlinear connections in synthetic differential geometry, *J. Pure Appl. Alg.* 131 (1998), 49–77; MR 99i:53011.

[908] H. Nishimura, Differential forms in synthetic differential supergeometry, *Internat. J. Theor. Phys.* 38 (1999), 653–663; MR 00h:58005.

[909] H. Nishimura, Infinitesimal calculus of variations, *Internat. J. Theor. Phys.* 38 (1999), 1771–1782; MR 00e:58028.

[910] H. Nishimura, General Jacobi identity revisited, *Internat. J. Theor. Phys.* 38 (1999), 2163–2174; MR 00h:58006.

[911] S. C. Nistor and I. Tofan, On the category $\mathbf{ab}(\mathcal{E})$ (Romanian), *Bul. Inst. Politehn. Iaşi* (I) 31 (1985), 205–207; MR 87g:18001.

[912] S. C. Nistor and I. Tofan, A functorial characterization of categories of modules in an ω-topos (Romanian), in *Proc. Conf. on Algebra, Timişoara 1986* (Univ. of Timişoara, 1987), 66–69; MR 89m:18004.

[913] A. Obtułowicz, The logic of categories of partial functions and its applications, *Dissertationes Math.* (*Rozprawy Mat.*) 241 (1986); MR 88b:03106

[914] A. Obtułowicz, Rough sets and Heyting algebra valued sets, *Bull. Pol. Acad. Sci. Math.* 35 (1987), 667–671; MR 89e:03110.

[915] A. Obtułowicz, Categorical and algebraic aspects of Martin-Löf type theory, *Studia Logica* 48 (1989), 299–317; MR 91k:03156.

[916] J. van Oosten, The modified realizability topos, *J. Pure Appl. Alg.* 116 (1997), 273–289; MR 98a:03096.

[917] J. van Oosten, Extensional realizability, *Ann. Pure Appl. Logic* 84 (1997), 317–349.

[918] J. van Oosten, Fibrations and calculi of fractions, *J. Pure Appl. Alg.* 146 (2000), 77–102; MR 00k:18008.

[919] J. van Oosten and A. K. Simpson, Axioms and (counter)examples in synthetic domain theory, *Ann. Pure Appl. Logic* 104 (2000), 233–278; MR 02f:68105.

[920] G. Osius, Categorical set theory: a characterization of the category of sets, *J. Pure Appl. Alg.* 4 (1974), 79–119; MR 51/643.

[921] G. Osius, The internal and external aspects of logic and set theory in elementary topoi, *Cahiers Top. Géom. Diff.* 15 (1974), 157–180; MR 52/7896.

[922] G. Osius, Logical and set-theoretical tools in elementary topoi, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 297–346; MR 52/7897.

[923] G. Osius, A note on Kripke–Joyal semantics for the internal language of topoi, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 349–354; MR 52/7898.

[924] E. Palmgren, Constructive sheaf semantics, *Math. Logic Quarterly* 43 (1997), 321–327; MR 98g:03151.

[925] E. Palmgren, A sheaf-theoretic foundation for nonstandard analysis, *Ann. Pure Appl. Logic* 85 (1997), 69–86; MR 97m:03098.

[926] E. Palmgren, Real numbers in the topos of sheaves over the category of filters, *J. Pure Appl. Alg.* 160 (2001), 275–284.

[927] R. Paré, Colimits in topoi, *Bull. Amer. Math. Soc.* 80 (1974), 556–561; MR 48/11245; also published in *Rev. Roum. Math. Pures Appl.* 22 (1977), 335–340; MR 56/442.

[928] R. Paré, Indexed categories and generated topologies, *J. Pure Appl. Alg.* 19 (1980), 385–400; MR 82c:18008.

[929] R. Paré, A topos with no geometric morphism to any Boolean one, *Math. Proc. Camb. Philos. Soc.* 97 (1985), 397–398; MR 86d:18003.

[930] R. Paré, Some applications of categorical model theory, in *Categories in computer science and logic*, Contemp. Math. vol. 92 (Amer. Math. Soc., 1989), 325–339; MR 91h:03048.

[931] R. Paré, Simply connected limits, *Canad. J. Math.* 42 (1990), 731–746; MR 91j:18004.

[932] R. Paré and L. Roman, Dinatural numbers, *J. Pure Appl. Alg.* 128 (1998), 33–92; MR 99f:18003.

[933] R. Paré and D. Schumacher, Abstract families and the adjoint functor theorems, in *Indexed categories and their applications*, Lecture Notes in Math. vol. 661 (Springer–Verlag, 1978), 1–125; MR 80b:18004.

[934] J. Paseka, T_2-separation axioms on frames, *Acta Univ. Carolin. Math. Phys.* 28 (1987), 95–98; MR 89a:54025.

[935] J. Paseka, Products in the category of locales: which properties are preserved?, *Discrete Math.* 108 (1992), 63–73; MR 94d:18006.

[936] J. Paseka and P. Sekanina, A note on extremally disconnected frames, *Acta Univ. Carolin. Math. Phys.* 31 (1990), 75–84; MR 92f:54035.

[937] J. Paseka and B. Šmarda, T_2-frames and almost compact frames, *Czech Math. J.* 42 (117) (1992), 385–402; MR 93i:18006.

[938] B. A. Pasynkov, Partial topological products, *Trans. Moscow Math. Soc.* 13 (1965), 153–272; MR 29/6469.

[939] D. Pataraia, Internal categories in a left exact cosimplicial category, *Georgian Math. J.* 4 (1997), 533–556; MR 98h:18002.

[940] V. Patryshev, On the Grothendieck topologies in the toposes of presheaves, *Cahiers Top. Géom. Diff.* 25 (1984), 207–215; MR 87a:18008.

[941] D. Pavlović, On completeness and cocompleteness in and around small categories, *Ann. Pure Appl. Logic* 74 (1995), 121–152; MR 96h:03106.

[942] M. C. Pedicchio, On Grothendieck quasi-topoi, in *Recent developments of general topology and its applications*, Math. Research vol. 67 (Akademie–Verlag, 1992), 252–255; MR 94c:18005.

[943] M. C. Pedicchio and F. Borceux, A remark on categories of separated objects, *Rend. Circ. Mat. Palermo Suppl.* (II) 24 (1990), 475–484; MR 92d:18007.

[944] M. C. Pedicchio and J. Rosický, Comparing coequalizer and exact completions, *Theory Appl. Categ.* 6 (1999), 77–82; MR 01b:18004.

[945] M. C. Pedicchio and W. Tholen, Localizations of locally presentable categories and exact topologies, *J. Pure Appl. Alg.* 64 (1990), 293–301; MR 91k:18007.

[946] M. C. Pedicchio and J. J. C. Vermeulen, A remark on the 'local character' condition, *Ann. Soc. Sci. Bruxelles* (I) 103 (1989), 73–85; MR 91j:18017.

[947] C. Pellegrino, Monomorfismi regolari e monomorfismi parziali: loro classificatore per una classe di categorie comma, *Atti Sem. Mat. Fis. Univ. Modena* 24 (1975), 300–305; MR 54/7572.

[948] C. Pellegrino, Construction of the subobject classifier for a class of comma categories (Italian), *Boll. Un. Mat. Ital.* B 16 (1979), 533–542; MR 82j:18003.

[949] J. W. Pelletier, Applications of the dual functor in Banach spaces, in *Mathematical applications of category theory*, Contemp. Math. vol. 30 (Amer. Math. Soc., 1984), 277–307; MR 85m:46071.

[950] J. W. Pelletier, Locales in functional analysis, *J. Pure Appl. Alg.* 70 (1991), 133–145; MR 92b:46121.

[951] J. W. Pelletier and R. D. Rosebrugh, The category of Banach spaces in sheaves, *Cahiers Top. Géom. Diff.* 20 (1979), 353–372; MR 81c:46055.

[952] J. W. Pelletier and R. D. Rosebrugh, On subcategories of Banach spaces in sheaves, in *Categorical aspects of topology and analysis*, Lecture Notes in Math. vol. 915 (Springer–Verlag, 1982), 260–274; MR 84k:46051.

[953] A. M. Penk, Two forms of the axiom of choice for an elementary topos, *J. Symbolic Logic* 40 (1975), 197–212; MR 51/5698.

[954] J. Penon, Quasitopos, *C. R. Acad. Sci. Paris* 276 (1973), A237–240; MR 48/11246.

[955] J. Penon, Catégories localement internes, *C. R. Acad. Sci. Paris* 278 (1974), A1577–1580; MR 51/10435.

[956] J. Penon, Sur les quasi-topos, *Cahiers Top. Géom. Diff.* 18 (1977), 181–218; MR 58/848.

[957] J. Penon, Infinitésimaux et intuitionnisme, *Cahiers Top. Géom. Diff.* 22 (1981), 67–72; MR 82m:51021.

[958] J. Penon, Logique de la géométrie algébrique, *Cahiers Top. Géom. Diff. Catégoriques* 33 (1992), 277–278.

[959] J. Penon, Construction d'infinitésimaux, *Cahiers Top. Géom. Diff. Catégoriques* 35 (1994), 339–349; MR 95j:54015.

[960] I. Petrescu, Model-companions for logical functors, *Rev. Roum. Math. Pures Appl.* 23 (1978), 413–417; MR 80a:03090.

[961] I. Petrescu, Existential morphisms and existentially closed models of logical categories, *Z. Math. Logik Grundl. Math.* 27 (1981), 363–370; MR 82k:03104.

[962] H. P. Phan, Regular topos I (Vietnamese), *Tap chi Toan hoc* 14 (1986), 18–26; MR 88m:18008.

[963] W. K. S. Phoa, Relative computability in the effective topos, *Math. Proc. Camb. Philos. Soc.* 106 (1989), 419–422; MR 90h:18004.

[964] W. K. S. Phoa, Effective domains and intrinsic structure, in *Proc. Fifth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1990), 366–377.

[965] L. Picco Botta, Sui punti e sugli aperti di un topos, *Univ. Politecn. Torino Rend. Sem. Mat.* 34 (1976), 189–200; MR 55/8144.

[966] L. Picco Botta, Topoi connessi, *Ist. Lombardo Accad. Sci. Lett. Rend.* (A) 111 (1977), 151–162; MR 58/16839.

[967] A. M. Pitts, Fuzzy sets do not form a topos, *Fuzzy Sets Syst.* 8 (1982), 101–104; MR 84e:03081.

[968] A. M. Pitts, Amalgamation and interpolation in the category of Heyting algebras, *J. Pure Appl. Alg.* 29 (1983), 155–165; MR 85c:03024.

[969] A. M. Pitts, An application of open maps to categorical logic, *J. Pure Appl. Alg.* 29 (1983), 313–326; MR 85c:03025.

[970] A. M. Pitts, On product and change of base for toposes, *Cahiers Top. Géom. Diff.* 26 (1985), 43–67; MR 87a:18009.

[971] A. M. Pitts, Polymorphism is set theoretic, constructively, in *Category theory and computer science*, Lecture Notes in Computer Science vol. 283 (Springer–Verlag, 1987), 12–39; MR 89c:03022.

[972] A. M. Pitts, Interpolation and conceptual completeness for pretoposes via category theory, in *Mathematical logic and theoretical computer science*, Lecture Notes Pure Appl. Math. vol. 106 (Marcel Dekker, 1987), 301–327; MR 89g:03096.

[973] A. M. Pitts, Applications of sup-lattice enriched category theory to sheaf theory, *Proc. London Math. Soc.* (3) 57 (1988), 433–480; MR 89m:18002.

[974] A. M. Pitts, Conceptual completeness for first order intuitionistic logic: an application of categorical logic, *Ann. Pure Appl. Logic* 41 (1989), 33–81; MR 90c:03065.

[975] A. M. Pitts, Non-trivial power types can't be subtypes of polymorphic types, in *Proc. Fourth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1989), 6–13.

[976] A. M. Pitts, On an interpretation of second-order quantification in first-order intuitionistic propositional logic, *J. Symbolic Logic* 57 (1992), 33–52; MR 93h:03009.

[977] A. M. Pitts and P. Taylor, A note on Russell's paradox in locally Cartesian closed categories, *Studia Logica* 48 (1989), 377–387; MR 91k:03018.

[978] T. Plewe, Localic products of spaces, *Proc. London Math. Soc.* (3) 73 (1996), 642–678; MR 97d:54012.

[979] T. Plewe, Localic triquotient maps are effective descent maps, *Math. Proc. Camb. Philos. Soc.* 122 (1997), 17–43; MR 98b:54022.

[980] T. Plewe, Quotient maps of locales, *Appl. Categ. Struct.* 8 (2000), 17–44; MR 01j:18009.

[981] T. Plewe, Higher order dissolutions and Boolean coreflections of locales, *J. Pure Appl. Alg.* 154 (2000), 273–293; MR 01f:18010.

[982] T. Plewe, A. Pultr and A. Tozzi, Regular monomorphisms of Hausdorff frames, *Appl. Categ. Struct.* 9 (2001), 15–33; MR 01k:18001.

[983] D. Ponasse, Some remarks on the category $\mathbf{Fuz}(H)$ of M. Eytan, *Fuzzy Sets Syst.* 9 (1983), 199–204; MR 84e:03082.

[984] D. Ponasse, Categorical studies of fuzzy sets, *Fuzzy Sets Syst.* 28 (1988), 235–244; MR 89k:03070.

[985] N. Popescu, *Abelian categories with applications to rings and modules*, L.M.S. Monographs no. 3 (Academic Press, 1973); MR 49/5130.

[986] H. J. Porta and O. Wyler, On compact space objects in topoi, in *Categorical aspects of topology and analysis*, Lecture Notes in Math. vol. 915 (Springer–Verlag, 1992), 375–385; MR 83g:18011.

[987] A. J. Power and E. P. Robinson, A characterization of pie limits, *Math. Proc. Camb. Philos. Soc.* 110 (1991), 33–47; MR 92d:18004.

[988] A. Preller, An interpretation of Martin-Löf's theory of types in elementary topos theory, *Z. Math. Logik Grundl. Math.* 38 (1992), 213–240; MR 95c:03146.

[989] M. Proietti, Connections between partial maps categories and tripos theory, in *Category theory and computer science*, Lecture Notes in Computer Science vol. 283 (Springer–Verlag, 1987), 254–269; MR 89a:18007.

[990] D. A. Pronk, Etendues and stacks as bicategories of fractions, *Compositio Math.* 102 (1996), 243–303; MR 97d:18011.

[991] A. I. Provotar, Incompleteness in a topos, *Cybernet. Systems Anal.* 32 (1996), 876–884; MR 98h:03082.

[992] A. I. Provotar, Linear morphisms in a topos, *Cybernet. Systems Anal.* 33 (1997), 155–161; MR 98e:18004.

[993] A. Pultr and W. Tholen, Localic enrichments of categories, *Cahiers Top. Géom. Diff. Catégoriques* 41 (2000), 121–142; MR 01i:18016.

[994] A. G. Radu, Sur les topos de Giraud–Grothendieck, *Ann. Fac. Sci. Kinshasa (Zaire), Sect. Math.-Phys.* 2 (1976), 108–118; MR 54/10381.

[995] A. G. Radu, Quelques catégories cartésiennes remarquables, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 25 (1979), 5–36; MR 81g:18014.

[996] A. G. Radu, Topologii Grothendieck pe categorii regulate şi categorii Boole, *Stud. Cerc. Mat.* 31 (1979), 463–478; MR 81c:18015.

[997] A. G. Radu, *Teoria toposurilor* (2 volumes) (Editura Acad. R. S. România, 1981–2); MR 82j:18001 and 84g:18014.

[998] A. G. Radu and G. G. Radu, Existential and universal quantifiers in an elementary topos, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 27 (1981), 19–30; MR 82k:18003.

[999] G. G. Radu and A. G. Radu, Topologia dublei negaţii, *Stud. Cerc. Mat.* 33 (1981), 547–552; MR 83m:03074.

[1000] G. G. Radu and A. G. Radu, Internal adjunction, *Rev. Roum. Math. Pures Appl.* 27 (1982), 187–198; MR 83m:18005.

[1001] J. Z. Reichman, Semicontinuous real numbers in a topos, *J. Pure Appl. Alg.* 28 (1983), 81–91; MR 84j:18002.

[1002] J. Reiterman, The category of complete lattices as a category of algebras, *Alg. Universalis* 26 (1989), 367–379; MR 91b:18006.

[1003] J. Reiterman and W. Tholen, Effective descent maps of topological spaces, *Top. Appl.* 57 (1994), 53–69; MR 95f:18008.

[1004] G. E. Reyes, From sheaves to logic, in *Studies in algebraic logic*, M.A.A Studies in Math. vol. 9 (Math. Assoc. of America, 1974), 143–204; MR 50/13182.

[1005] G. E. Reyes, Sheaves and concepts: a model-theoretic introduction to Grothendieck topoi, *Cahiers Top. Géom. Diff.* 18 (1977), 105–137; MR 58/5184.

[1006] G. E. Reyes, Théorie des modèles et faisceaux, *Adv. Math.* 30 (1978), 156–170; MR 81f:03050.

[1007] G. E. Reyes, Cramer's rule in the Zariski topos, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 586–594; MR 81e:03063.

[1008] G. E. Reyes, Analyse dans les topos lisses, *Cahiers Top. Géom. Diff.* 22 (1981), 156–160; MR 83m:13018.

[1009] G. E. Reyes, A topos-theoretic approach to reference and modality, *Notre Dame J. Formal Logic* 32 (1991), 359–391; MR 92m:03039.

[1010] G. E. Reyes and N. V. Quê, Smooth functors and synthetic calculus, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 377–395; MR 85d:58006.

[1011] G. E. Reyes and G. C. Wraith, A note on tangent bundles in a category with a ring object, *Math. Scand.* 42 (1978), 53–63; MR 80f:18011.

[1012] G. E. Reyes and H. Zolfaghari, Topos-theoretic approaches to modality, in *Category theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 359–378; MR 93h:18009.

[1013] G. E. Reyes and H. Zolfaghari, Bi-Heyting algebras, toposes and modalities, *J. Philos. Logic* 25 (1996), 25–43; MR 97d:03020.

[1014] G. D. Reynolds, Alexandroff algebras and complete regularity, *Proc. Amer. Math. Soc.* 76 (1979), 322–326; MR 81d:54014.

[1015] G. D. Reynolds, On the spectrum of a real representable ring, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 595–611; MR 81c:18008.

[1016] M. D. Rice, Discrete uniform structures and quasitopoi, *Quaest. Math.* 7 (1984), 385–396; MR 86e:18007.

[1017] M. Richter, Limites in Kategorien von Relationalsystemen, *Z. Math. Logik Grundl. Math.* 17 (1971), 75–90; MR 43/4878.

[1018] E. P. Robinson, Stable theories of local rings, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 257–280; MR 85d:13040.

[1019] E. P. Robinson, The p-adic spectrum, *J. Pure Appl. Alg.* 40 (1986), 281–296; MR 88d:14002.

[1020] E. P. Robinson, The geometric theory of p-adic fields, *J. Algebra* 110 (1987), 158–172; MR 88k:12011.

[1021] E. P. Robinson, How complete is PER?, in *Proc. Fourth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1989), 106–111.

[1022] E. P. Robinson and G. Rosolini, Categories of partial maps, *Info. Comput.* 79 (1988), 95–130; MR 90b:18010a.

[1023] E. P. Robinson and G. Rosolini, Colimit completions and the effective topos, *J. Symbolic Logic* 55 (1990), 678–699; MR 91f:03130.
[1024] E. P. Robinson and G. Rosolini, Polymorphism, set theory and call-by-value, in *Proc. Fifth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1990), 12–18.
[1025] R. O. Robson, Model theory and spectra, *J. Pure Appl. Alg.* 63 (1990), 301–327; MR 91n:03039.
[1026] P. Rocos, D'un isomorphisme de deux catégories, *C. R. Acad. Sci. Paris* 282 (1976), A179–181; MR 54/2759.
[1027] J. R. Roisin, Introduction à une approche catégorique en théorie des modèles, *Bull. Soc. Math. Belg.* (B) 29 (1977), 161–174; MR 57/16068.
[1028] J. R. Roisin, On functorializing usual first-order model theory, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 612–622; MR 81f:03051.
[1029] L. Roman, Ultradiophantine categories, *Z. Math. Logik Grundl. Math.* 34 (1988), 289–295; MR 90d:03139.
[1030] L. Roman, Cartesian categories with natural numbers object, *J. Pure Appl. Alg.* 58 (1989), 267–278; MR 91a:18012.
[1031] J. E. Roos, Sur la distributivité des foncteurs $\varprojlim$ par rapport aux $\varinjlim$ dans les catégories de faisceaux (topos), *C. R. Acad. Sci. Paris* 259 (1964), 969–972, 1605–1608 and 1801–1804; MR 30/4816 and 32/5714–5.
[1032] R. D. Rosebrugh, Coequalizers in algebras for an internal type, in *Indexed categories and their applications*, Lecture Notes in Math. vol. 661 (Springer–Verlag, 1978), 243–260; MR 80m:03114.
[1033] R. D. Rosebrugh, On algebras defined by operations and equations in a topos, *J. Pure Appl. Alg.* 17 (1980), 203–221; MR 81d:18011.
[1034] R. D. Rosebrugh, On defining objects by recursion in a topos, *J. Pure Appl. Alg.* 20 (1981), 325–335; MR 82e:18006.
[1035] R. D. Rosebrugh, On endomorphisms of internal presheaf topoi, *Commun. Alg.* 9 (1981), 1901–1912; MR 83c:18004.
[1036] R. D. Rosebrugh, Bounded functors, finite limits and an application to injective topoi, *Cahiers Top. Géom. Diff.* 24 (1983), 267–278; MR 85h:18003.
[1037] R. D. Rosebrugh and R. J. Wood, Cofibrations in the bicategory of topoi, *J. Pure Appl. Alg.* 32 (1984), 71–94; MR 85j:18008.
[1038] R. D. Rosebrugh and R. J. Wood, Cofibrations II: left exact right actions and composition of gamuts, *J. Pure Appl. Alg.* 39 (1986), 283–300; MR 87e:18011.
[1039] R. D. Rosebrugh and R. J. Wood, Proarrows and cofibrations, *J. Pure Appl. Alg.* 53 (1988), 271–296; MR 89i:18005.
[1040] R. D. Rosebrugh and R. J. Wood, Gamuts and cofibrations, *Cahiers Top. Géom. Diff. Catégoriques* 31 (1990), 197–211; MR 92c:18007.
[1041] R. D. Rosebrugh and R. J. Wood, Pullback preserving functors, *J. Pure Appl. Alg.* 73 (1991), 73–90; MR 92g:18004.
[1042] R. D. Rosebrugh and R. J. Wood, Constructive complete distributivity, II, *Math. Proc. Camb. Philos. Soc.* 110 (1991), 245–249; MR 92j:06012.
[1043] R. D. Rosebrugh and R. J. Wood, Constructive complete distributivity, III, *Canad. Math. Bull.* 35 (1992), 537–547; MR 93k:06020.
[1044] R. D. Rosebrugh and R. J. Wood, Constructive complete distributivity, IV, *Appl. Categ. Struct.* 2 (1994), 119–144; MR 95e:06030.
[1045] R. D. Rosebrugh and R. J. Wood, An adjoint characterization of the category of sets, *Proc. Amer. Math. Soc.* 122 (1994), 409–413; MR 95a:18003.
[1046] R. D. Rosebrugh and R. J. Wood, Distributive adjoint strings, *Theory Appl. Categ.* 1 (1995), 119–145; MR 96h:18004.

[1047] K. I. Rosenthal, Étendues and categories with monic maps, *J. Pure Appl. Alg.* 22 (1981), 193–212; MR 83a:18013.

[1048] K. I. Rosenthal, Local equivalence relations, *Top. Appl.* 13 (1982), 167–176; MR 83c:54074.

[1049] K. I. Rosenthal, Quotient systems in Grothendieck topoi, *Cahiers Top. Géom. Diff.* 23 (1982), 425–438; MR 84m:18007.

[1050] K. I. Rosenthal, Sheaves and local equivalence relations, *Cahiers Top. Géom. Diff.* 25 (1984), 179–206; MR 87c:54012.

[1051] K. I. Rosenthal, Covering étendues and Freyd's theorem, *Proc. Amer. Math. Soc.* 99 (1987), 221–222; MR 88a:18005.

[1052] K. I. Rosenthal, *Quantales and their applications*, Pitman Res. Notes Math. no. 234 (Longman, 1990); MR 92e:06028.

[1053] J. Rosický, Categories of models of infinitary Horn theories, *Arch. Math. (Brno)* 14 (1978), 219–226; MR 80h:18004.

[1054] J. Rosický, Concrete categories and infinitary languages, *J. Pure Appl. Alg.* 22 (1981), 309–339; MR 82m:18004.

[1055] J. Rosický, A note on exponentiation in regular locales, *Arch. Math. (Brno)* 22 (1986), 157–158; MR 87m:18008.

[1056] J. Rosický, A categorical characterization of sets among classes, *Arch. Math. (Brno)* 23 (1987), 117–120; MR 89a:03089.

[1057] J. Rosický, Elementary categories, *Arch. Math. (Basel)* 52 (1989), 284–288; MR 90i:03043.

[1058] J. Rosický, Models of Horn theories revisited, *J. Pure Appl. Alg.* 92 (1994), 185–190; MR 94m:03057.

[1059] J. Rosický, More on directed colimits of models, *Appl. Categ. Struct.* 2 (1994), 71–76; MR 95j:03072.

[1060] J. Rosický, Cartesian closed exact completions, *J. Pure Appl. Alg.* 142 (1999), 261–270; MR 00h:18010.

[1061] J. Rosický and B. Šmarda, T_1-locales, *Math. Proc. Camb. Philos. Soc.* 98 (1985), 81–86; MR 86h:54027.

[1062] J. Rosický and W. Tholen, Accessibility and the solution set condition, *J. Pure Appl. Alg.* 98 (1995), 189–208; MR 95m:18001.

[1063] J. Rosický, V. Trnková and J. Adámek, Unexpected properties of locally presentable categories, *Alg. Universalis* 27 (1990), 153–170; MR 91a:18001.

[1064] G. Rosolini, Semireticoli e cornici in un topos, *Riv. Mat. Univ. Parma* (4) 4 (1978), 369–384; MR 81i:03013.

[1065] G. Rosolini, Estensioni mediante cornici in un topos, *Riv. Mat. Univ. Parma* (4) 7 (1981), 225–232; MR 83m:03075.

[1066] G. Rosolini, Domains and dominical categories, *Riv. Mat. Univ. Parma* (4) 11 (1986), 387–397; MR 87m:18010.

[1067] G. Rosolini, About modest sets, *Internat. J. Found. Comp. Sci.* 1 (1990), 341–353; MR 92b:68047.

[1068] C. Rousseau, Topos theory and complex analysis, *J. Pure Appl. Alg.* 10 (1977), 299–313; MR 57/413.

[1069] C. Rousseau, Topos theory and complex analysis, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer-Verlag, 1979), 623–659; MR 81d:32014.

[1070] C. Rousseau, Complex structures on topoi, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 196–210; MR 81h:53001.

[1071] C. Rousseau, Nombres réels et complexes dans les topos spatiaux, *Ann. Sci. Math. Québec* 3 (1979), 143–154; MR 80c:18004.

[1072] C. Rousseau, Formes normales des matrices rectangulaires dans les topos et applications, *Ann. Sci. Math. Québec* 5 (1981), 81–85; MR 82k:18004.

[1073] C. Rousseau, Spectral decomposition theorem for real symmetric matrices in topoi and applications, *J. Pure Appl. Alg.* 38 (1985), 91–102; MR 87b:18005.
[1074] D. Rovella, Sul lemma di Nakayama in un topos elementare, *Ist. Lombardo Accad. Sci. Lett. Rend.* (A) 110 (1976), 420–429; MR 58/11072.
[1075] K. A. Rowe, Images in topoi, *Canad. Math. Bull.* 20 (1977), 471–478; MR 57/16375.
[1076] W. B. G. Ruitenburg, On the period of sequences $(A^n(p))$ in intuitionistic propositional calculus, *J. Symbolic Logic* 49 (1984), 892–899; MR 86b:03016.
[1077] I. B. Rumbos, The Lambek sheaf re-examined, *Commun. Alg.* 19 (1991), 2291–2301; corrigenda *ibid.* 20 (1992), 623; MR 92m:18018.
[1078] G. Sambin, Intuitionistic formal spaces – a first communication, in *Mathematical logic and its applications* (Plenum Press, 1987), 187–204; MR 89f:03056.
[1079] G. Sambin, Intuitionistic formal spaces and their neighbourhood, in *Logic Colloquium '88*, Studies in Logic and the Foundations of Math. vol. 127 (North–Holland, 1989), 261–285; MR 90i:03069.
[1080] G. Sambin, Pretopologies and completeness proofs, *J. Symbolic Logic* 60 (1995), 861–878; MR 96m:03015.
[1081] G. Sambin, Formal topology and domains, *Elect. Notes Theor. Comp. Sci.* 35 (2000), 14pp; MR 01k:54003.
[1082] H. Sato, E-ccc: between ccc and topoi, *Theoret. Comput. Sci.* 64 (1989), 55–66; MR 90e:68068.
[1083] A. Ščedrov, Consistency and independence results in intuitionistic set theory, in *Constructive mathematics*, Lecture Notes in Math. vol. 873 (Springer–Verlag, 1981), 54–86; MR 83a:03062.
[1084] A. Ščedrov, Independence of the Fan theorem in the presence of continuity principles, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 435–442; MR 85b:03106.
[1085] A. Ščedrov, *Forcing and classifying topoi*, Memoirs Amer. Math. Soc. no. 295 (1984); MR 86d:03057.
[1086] A. Ščedrov, Differential equations in constructive analysis and in the recursive realizability topos, *J. Pure Appl. Alg.* 33 (1984), 69–80; MR 85h:03063.
[1087] A. Ščedrov, On some non-classical extensions of second-order intuitionistic propositional calculus, *Ann. Pure Appl. Logic* 27 (1984), 155–164; MR 86f:03022.
[1088] A. Scedrov, Diagonalization of continuous matrices as a representation of intuitionistic reals, *Ann. Pure Appl. Logic* 30 (1986), 201–206; MR 87k:03078.
[1089] A. Scedrov, Some aspects of categorical semantics: sheaves and glueing, in *Logic Colloquium '85*, Studies in Logic and the Foundations of Math. vol. 122 (North–Holland, 1987), 281–301; MR 88g:03088.
[1090] A. Scedrov, Kleene computable functions and the higher order existence property, *J. Pure Appl. Alg.* 52 (1988), 313–320; MR 89g:18008.
[1091] A. Ščedrov and P. J. Scott, A note on the Friedman slash and Freyd covers, in *The L.E.J. Brouwer centenary symposium*, Studies in Logic and the Foundations of Math. vol. 110 (North–Holland, 1982), 443–452; MR 85d:03118.
[1092] A. Scedrov and P. Scowcroft, Decompositions of finitely generated modules over $C(X)$: sheaf semantics and a decision procedure, *Math. Proc. Camb. Philos. Soc.* 103 (1988), 257–268; MR 89c:46095.
[1093] A. Schalk, Domains arising as algebras for powerspace constructions, *J. Pure Appl. Alg.* 89 (1993), 305–328; MR 95a:18004.
[1094] G. Schlitt, The Lindelöf Tychonoff Theorem and choice principles, *Math. Proc. Camb. Philos. Soc.* 110 (1991), 57–65; MR 92b:03056.

[1095] G. Schlitt, Sheaves of abelian groups and the quotients A^{**}/A, *J. Algebra* 158 (1993), 50–60; MR 94e:20072.

[1096] D. Schlomiuk, An elementary theory of the category of topological spaces, *Trans. Amer. Math. Soc.* 149 (1970), 259–278; MR 41/3559.

[1097] D. Schlomiuk, Topos di Grothendieck e topos di Lawvere e Tierney, *Rend. Mat.* (VI) 7 (1974), 513–553; MR 51/3252.

[1098] M. Schönfinkel, Über die Bausteine der mathematischen Logik, *Math. Annalen* 92 (1924), 305–316.

[1099] D. Schumacher, Absolutely free algebras in a topos containing an infinite object, *Canad. Math. Bull.* 19 (1976), 323–328; MR 55/12790.

[1100] D. Schumacher, Restricted internal choice in a topos, *Commun. Alg.* 12 (1984), 81–101; MR 85c:18004.

[1101] D. Schumacher and R. Street, Some parametrized categorical concepts, *Commun. Alg.* 16 (1988), 2313–2347; MR 89k:18001.

[1102] D. S. Scott, Continuous lattices, in *Toposes, algebraic geometry and logic*, Lecture Notes in Math. vol. 274 (Springer–Verlag, 1972), 97–136; MR 53/7879.

[1103] D. S. Scott, Identity and existence in intuitionistic logic, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 660–696; MR 81a:03060.

[1104] D. S. Scott, Relating theories of the λ-calculus, in *To H. B. Curry: Essays on combinatory logic, lambda-calculus and formalism* (Academic Press, 1980), 403–450; MR 82d:03025.

[1105] R. A. G. Seely, Weak adjointness in proof theory, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 697–701; MR 81b:03063.

[1106] R. A. G. Seely, Hyperdoctrines, natural deduction and the Beck condition, *Z. Math. Logik Grundl. Math.* 29 (1983), 505–542; MR 86c:03054.

[1107] R. A. G. Seely, Locally Cartesian closed categories and type theory, *Math. Proc. Camb. Philos. Soc.* 95 (1984), 33–48; MR 86b:18008.

[1108] V. M. Sergeev, On the topos structure on categories of topos-valued functors (Russian), *Cestnik Moskov. Univ.* I (1980), 38–42; MR 82a:18003.

[1109] W. Sierpiński, L'axiome de M. Zermelo et son rôle dans la théorie des ensembles et l'analyse, *Bull. Internat. Acad. Sci. Cracovie* A (1918), 97–152.

[1110] H. Simmons, A framework for topology, in *Logic Colloquium 1977*, Studies in Logic and the Foundations of Math. vol. 96 (North-Holland, 1978), 239–251; MR 80b:06008.

[1111] H. Simmons, An algebraic version of Cantor–Bendixson analysis, in *Categorical aspects of topology and analysis*, Lecture Notes in Math. vol. 915 (Springer–Verlag, 1982), 310–323; MR 83i:06022.

[1112] H. Simmons, The glueing construction and lax limits, *Math. Struct. Comput. Sci.* 4 (1994), 393–431; MR 96d:18002.

[1113] A. K. Simpson, Computational adequacy in an elementary topos, in *Computer science logic*, Lecture Notes in Computer Science vol. 1584 (Springer–Verlag, 1999), 323–342; MR 00g:68093.

[1114] J. D.H. Smith, *Mal'cev varieties*, Lecture Notes in Math. vol. 554 (Springer–Verlag, 1976); MR 55/5499.

[1115] I. Sols, Bon ordre dans l'objet des nombres naturels d'un topos booléen, *C. R. Acad. Sci. Paris* 281 (1975), A601–603; MR 52/96.

[1116] I. Sols, Programming in topoi, *Cahiers Top. Géom. Diff.* 16 (1976), 312–319.

[1117] I. Sols, Logic, foundations, geometry and automata in the area of toposes (Spanish), in *Coloquio sobre Logica Simbolica, Madrid 1975* (Universidad Complutense, 1976), 159–176; MR 56/15419.

[1118] A. C. Ştefănescu, The category of abelian groups in a topos, *Bull. Soc. Sci. Math. Roum.* 25 (73) (1981), 179–189; MR 83c:18005.
[1119] A. C. Ştefănescu, Structura obiectului de numere integri şi de numere raţionale intr-un topos Peano, *Studii Cerc. Mat.* 34 (1982), 568–579; MR 85a:18006a.
[1120] A. C. Ştefănescu, Grupul aditiv al numerelor reale Cauchy intr-un topos Peano, *Studii Cerc. Mat.* 35 (1983), 220–232; MR 85a:18006b.
[1121] T. C. Stevens, Connectedness of complete metric groups, *Colloq. Math.* 50 (1986), 233–240; MR 88h:54046.
[1122] L. N. Stout, Topological space objects in a topos II: $\mathcal{E}$-completeness and $\mathcal{E}$-cocompleteness, *Manuscripta Math.* 17 (1975), 1–14; MR 54/2751.
[1123] L. N. Stout, Topological properties of the real numbers object in a topos, *Cahiers Top. Géom. Diff.* 17 (1976), 295–326; MR 54/10015.
[1124] L. N. Stout, Quels sont les espaces topologiques dans les topos?, *Ann. Sci. Math. Québec* 2 (1978), 123–141; MR 80d:18007.
[1125] L. N. Stout, Laminations, or how to build a quantum-logic-valued model of set theory, *Manuscripta Math.* 28 (1979), 379–403; MR 81i:03099.
[1126] L. N. Stout, Topoi and categories of fuzzy sets, *Fuzzy Sets Syst.* 12 (1984), 169–184; MR 85i:18008.
[1127] L. N. Stout, Dedekind finiteness in topoi, *J. Pure Appl. Alg.* 49 (1987), 219–225; MR 88m:18009.
[1128] L. N. Stout, A survey of fuzzy set theory and topos theory, *Fuzzy Sets Syst.* 42 (1991), 3–14; MR 92h:03087.
[1129] R. Street, Elementary cosmoi I, in *Proc. Sydney Category Seminar 1973*, Lecture Notes in Math. vol. 420 (Springer–Verlag, 1974), 134–180; MR 50/7290.
[1130] R. Street, Limits indexed by category-valued 2-functors, *J. Pure Appl. Alg.* 8 (1976), 350–379; MR 53/5695.
[1131] R. Street, Cosmoi of internal categories, *Trans. Amer. Math. Soc.* 258 (1980), 271–318; MR 82a:18007.
[1132] R. Street, Fibrations in bicategories, *Cahiers Top. Géom. Diff.* 21 (1980), 111–160; MR 81f:18028.
[1133] R. Street, Conspectus of variable categories, *J. Pure Appl. Alg.* 21 (1981), 307–338; MR 85i:18006.
[1134] R. Street, Notions of topos, *Bull. Austral. Math. Soc.* 23 (1981), 199–208; MR 83a:18014.
[1135] R. Street, Two-dimensional sheaf theory, *J. Pure Appl. Alg.* 23 (1982), 251–270; MR 83d:18014.
[1136] R. Street, Characterization of bicategories of stacks, in *Category theory*, Lecture Notes in Math. vol. 962 (Springer–Verlag, 1982), 282–291; MR 84d:18006.
[1137] R. Street, The family approach to total cocompleteness and toposes, *Trans. Amer. Math. Soc.* 284 (1984), 355–369; MR 85f:18003.
[1138] R. Street, The petit topos of globular sets, *J. Pure Appl. Alg.* 154 (2000), 299–315; MR 01f:18008.
[1139] R. Street and R. F. C. Walters, The comprehensive factorization of a functor, *Bull. Amer. Math. Soc.* 79 (1973), 936–941; MR 49/10753.
[1140] R. Street and R. F. C. Walters, Yoneda structures on 2-categories, *J. Algebra* 50 (1978), 350–379; MR 57/3214.
[1141] J. R. Strooker, *Introduction to categories, homological algebra and sheaf cohomology* (Cambridge Univ. Press, 1978); MR 58/16815.
[1142] S.-H. Sun, A localic approach to minimal prime spectra, *Math. Proc. Camb. Philos. Soc.* 103 (1988), 47–53; MR 88j:06035.
[1143] M. E. Szabo, *Algebra of proofs*, Studies in Logic and the Foundations of Math. vol. 88 (North-Holland, 1978); MR 80b:03097.

[1144] M. E. Szabo, Categorical De Morgan laws, *Alg. Universalis* 12 (1981), 93–102; MR 83c:03057.

[1145] M. E. Szabo, The continuous realizability of entailment, *Z. Math. Logik Grundl. Math.* 29 (1983), 219–233; MR 85e:03050.

[1146] M. E. Szabo, Variable truth, *Z. Math. Logik Grundl. Math.* 30 (1984), 401–414; MR 86e:03030.

[1147] C. Szasz, Das Objekt 'ganze Zahlen' in einem elementaren Topos, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 31 (1985), 88–89; MR 88a:18006.

[1148] G. Szeto, The sheaf representation of near-rings and its applications, *Commun. Alg.* 5 (1977), 773–782; MR 55/10534.

[1149] G. Szeto, On the sheaf representation of a biregular near-ring, *Canad. Math. Bull.* 20 (1977), 495–500; MR 57/9762.

[1150] G. Takeuti, Boolean valued analysis, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 714–731; MR 81c:03045.

[1151] J. Tavakoli, Elements of free modules in topoi, *Commun. Alg.* 10 (1982), 171–201; MR 84a:18006.

[1152] J. Tavakoli, Normal forms of matrices in topoi, *Cahiers Top. Géom. Diff.* 23 (1982), 317–340; MR 84c:18009.

[1153] J. Tavakoli, A characterization of geometric fields in a topos, *Commun. Alg.* 13 (1984), 101–111; MR 86b:18005.

[1154] J. Tavakoli, Tensor products of modules in a topos, *Bull. Iran. Math. Soc.* 11 (1984), 23–32; MR 86i:18004.

[1155] J. Tavakoli, On products of matrices in a topos, *J. Austral. Math. Soc.* A 38 (1985), 416–420; MR 86e:18005.

[1156] J. Tavakoli, On topologies in a topos, *Bull. Austral. Math. Soc.* 41 (1990), 387–392; MR 91g:18003.

[1157] J. Tavakoli, A note on the sheaf functor, *Bull. Iran. Math. Soc.* 18 (1992), 1–7; MR 93d:18005.

[1158] J. Tavakoli, Locally free vector spaces in a topos, *Commun. Alg.* 24 (1996), 1199–1205; MR 96m:18010.

[1159] P. Taylor, The fixed point property in synthetic domain theory, in *Proc. Sixth IEEE Symp. on Logic in Computer Science* (IEEE Computer Soc. Press, 1991), 152–160.

[1160] P. Taylor, Intuitionistic sets and ordinals, *J. Symbolic Logic* 61 (1996), 705–744; MR 97j:03102.

[1161] P. Taylor, *Practical foundations of mathematics*, Cambridge Studies in Advanced Math. no. 59 (Cambridge Univ. Press, 1999); MR 00h:03002.

[1162] P. Taylor, Geometric and higher order logic in terms of abstract Stone duality, *Theory Appl. Categ.* 7 (2000), 284–338; MR 02b:03137.

[1163] B. R. Tennison, *Sheaf theory*, L.M.S. Lecture Notes Series no. 20 (Cambridge Univ. Press, 1975); MR 53/8192.

[1164] M.-F. Thibault, Pre-recursive categories, *J. Pure Appl. Alg.* 24 (1982), 79–93; MR 83f:03040.

[1165] R. W. Thomason, The finite stable homotopy type of some topoi, *J. Pure Appl. Alg.* 47 (1987), 89–104; MR 88k:14010.

[1166] M. Tierney, Sheaf theory and the continuum hypothesis, in *Toposes, algebraic geometry and logic*, Lecture Notes in Math. vol. 274 (Springer–Verlag, 1972), 13–42; MR 51/10088.

[1167] M. Tierney, Axiomatic sheaf theory: some constructions and applications, in *Proc. C.I.M.E. Summer School on Categories and Commutative Algebra, Varenna 1971* (Edizioni Cremonese, 1973), 249–326; MR 50/7277.

[1168] M. Tierney, On the spectrum of a ringed topos, in *Algebra, topology and category theory: a collection of papers in honor of Samuel Eilenberg* (Academic Press, 1976), 189–210; MR 53/13356.

[1169] M. Tierney, Forcing topologies and classifying topoi, in *Algebra, topology and category theory: a collection of papers in honor of Samuel Eilenberg*, Academic Press (1976), 211–219; MR 53/13357.
[1170] I. Tofan, Sur les catégories internes, *An. Ştiinţ. Univ. din Iaşi Secţ. IA Mat. (N.S.)* 26 (1980), 267–269; MR 82e:18007.
[1171] I. Tofan, Triples in elementary topoi (Romanian), *Bul. Inst. Politehn. Iaşi Sect. 1* (1985), 209–211; MR 87d:18005.
[1172] H. Toth, Categorical properties of f-set theory, *Fuzzy Sets Syst.* 33 (1989), 99–109; MR 90m:03096.
[1173] C. F. Townsend, Localic Priestley duality, *J. Pure Appl. Alg.* 116 (1997), 323–335; MR 98c:06012.
[1174] M. S. Tsalenko, The structure of correspondence categories, *Soviet Math. Dokl.* 18 (1977), 1498–1502.
[1175] H. Tsukada, Category theory not based upon set theory, *Sci. Papers Coll. Gen. Ed. Univ. Tokyo* 31 (1981), 1–24; MR 83d:18001.
[1176] M. Turgi, A sheaf-theoretic interpretation of the Kuroš theorem, in *Abelian group theory*, Lecture Notes in Math. vol. 616 (Springer–Verlag, 1977), 173–196; MR 57/308.
[1177] T. Uesu, Intuitionistic theories and toposes, in *Logic Symposia, Hakone 1979, 1980*, Lecture Notes in Math. vol. 891 (Springer–Verlag, 1981), 323–358; MR 84f:03057.
[1178] K.-H. Ulbrich, On the correspondence between gerbes and bouquets, *Math. Proc. Camb. Philos. Soc.* 108 (1990), 1–5; MR 91j:18010.
[1179] K.-H. Ulbrich, On cocycle bitorsors and gerbes over a Grothedieck topos, *Math. Proc. Camb. Philos. Soc.* 110 (1991), 49–55; MR 92g:18006.
[1180] F. Ulmer, Locally α-presentable and locally α-generated categories, in *Reports of the Midwest category seminar V*, Lecture Notes in Math. vol. 195 (Springer–Verlag, 1971), 230–247; MR 45/1990.
[1181] F. Ulmer, On the existence and exactness of the associated sheaf functor, *J. Pure Appl. Alg.* 3 (1973), 295–306; MR 51/10431.
[1182] V. Vajner, The category of complete Boolean algebras is not an intersection of reflective subcategories of the category of frames, *Appl. Categ. Struct.* 1 (1993), 191–195; MR 94k:18001.
[1183] S. Valentini, Points and co-points in formal topology, *Boll. Un. Mat. Ital.* (VII) 7–A (1993), 7–19; MR 93k:54025.
[1184] S. Valentini, A completeness theorem for formal topologies, in *Logic and Algebra*, Lecture Notes Pure Appl. Math. vol. 180 (Marcel Dekker, 1996), 689–702; MR 97g:06009.
[1185] P. Vámos, Sheaf-theoretical methods in the solution of Kaplansky's problem, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 732–738; MR 81h:13008.
[1186] G. Van Den Bossche, Algèbres de Heyting complètes étalées et faisceaux, *Ann. Soc. Sci. Bruxelles* 91 (1977), 107–116; MR 58/5854.
[1187] G. Van De Wauw–De Kinder, Arithmétique de premier ordre dans les topos, *C. R. Acad. Sci. Paris* 280 (1975), A1579–1582; MR 52/8217.
[1188] D. H. Van Osdol, Sheaves in regular categories, in *Exact categories and categories of sheaves*, Lecture Notes in Math. vol. 236 (Springer–Verlag, 1971), 233–239.
[1189] D. H. Van Osdol, Coalgebras, sheaves and cohomology, *Proc. Amer. Math. Soc.* 33 (1972), 257–263; MR 45/3517.
[1190] D. H. Van Osdol, Homological algebra in topoi, *Proc. Amer. Math. Soc.* 50 (1975), 52–54; MR 51/5714.

[1191] D. H. Van Osdol, Simplicial homotopy in an exact category, *Amer. J. Math.* 99 (1977), 1193–1204; MR 57/17649.

[1192] D. H. Van Osdol, A Borel topos, *Cahiers Top. Géom. Diff.* 22 (1981), 123–128; MR 84h:18013.

[1193] V. L. Vasyukov, Quantum logic in topoi (Russian), in *Investigations in nonclassical logics* (Russian), Nauka (1989), 333–348; MR 91k:03148.

[1194] V. L. Vasyukov, Implicative logics in categories, *Bull. Sect. Logic Univ. Łódź* 26 (1997), 188–192.

[1195] B. Veit Riccioli, Il forcing come principio logico per la costruzione dei fasci, I, *Rend. Mat.* (VI) 11 (1978), 329–353; MR 80b:03109a.

[1196] B. Veit Riccioli, Il forcing come principio logico per la costruzione dei fasci, II, *Rend. Mat.* (VI) 11 (1978), 601–625; MR 80b:03109b.

[1197] B. Veit, A proof of the associated sheaf theorem by means of categorical logic, *J. Symbolic Logic* 46 (1981), 45–55; MR 82d:03103.

[1198] B. Veit, Convergence in some ringed toposes, *Cahiers Top. Géom. Diff. Catégoriques* 31 (1990), 245–274; MR 92b:18004.

[1199] B. Veit, Sheaves, localizations and unstable extensions: some counterexamples, *J. Algebra* 140 (1991), 370–391; MR 92i:18005.

[1200] J. J. C. Vermeulen, Some constructive results related to compactness and the strong Hausdorff property for locales, in *Category Theory*, Lecture Notes in Math. vol. 1488 (Springer–Verlag, 1991), 401–409; MR 93g:54033.

[1201] J. J. C. Vermeulen, A note on iterative arguments in a topos, *Math. Proc. Camb. Philos. Soc.* 110 (1991), 57–62; MR 92h:18003.

[1202] J. J. C. Vermeulen, Weak compactness in constructive spaces, *Math. Proc. Camb. Philos. Soc.* 111 (1992), 63–74; MR 92i:54028.

[1203] J. J. C. Vermeulen, Proper maps of locales, *J. Pure Appl. Alg.* 92 (1994), 79–107; MR 95d:54009.

[1204] J. J. C. Vermeulen, A note on stably closed maps of locales, *J. Pure Appl. Alg.* 157 (2001), 335–339; MR 01m:06019.

[1205] S. J. Vickers, *Topology via logic*, Cambridge Tracts in Theoret. Comp. Sci. no. 5 (Cambridge Univ. Press, 1989); MR 90j:03110.

[1206] S. J. Vickers, Geometric theories and databases, in *Applications of categories in computer science*, L.M.S. Lecture Notes Series no. 177 (1992), 288–314; MR 93j:68139.

[1207] S. J. Vickers, Information systems for continuous posets, *Theoret. Comput. Sci.* 114 (1993), 201–229; MR 95c:68148.

[1208] S. J. Vickers, Toposes pour les vraiment nuls, in *Advances in theory and formal methods of computing* (Imperial College Press, 1996), 1–12.

[1209] S. J. Vickers, Constructive points of powerlocales, *Math. Proc. Camb. Philos. Soc.* 122 (1997), 207–222; MR 98h:18003.

[1210] S. J. Vickers, Topical categories of domains, *Math. Struct. Comput. Sci.* 9 (1999), 569–616; MR 00h:18007.

[1211] E. M. Vitale, Characterization of monadic categories over **Set**, *Cahiers Top. Géom. Diff. Catégoriques* 35 (1994), 351–358; MR 95m:18003.

[1212] E. M. Vitale, Localizations of algebraic categories, *J. Pure Appl. Alg.* 108 (1996), 315–320; MR 96m:18013.

[1213] E. M. Vitale, Localizations of algebraic categories, II, *J. Pure Appl. Alg.* 133 (1998), 317–326; MR 99h:18007.

[1214] E. M. Vitale, Essential localizations and infinitary exact completion, *Theory Appl. Categ.* 8 (2001), 465–480; MR 02f:18005.

[1215] H. Volger, Completeness theorem for logical categories, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 51–86; MR 51/12983.

[1216] H. Volger, Logical categories, semantical categories and topoi, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 87–100; MR 51/12984.

[1217] H. Volger, Ultrafilters, ultrapowers and finiteness in a topos, *J. Pure Appl. Alg.* 6 (1975), 345–356; MR 52/84.

[1218] H. Volger, Preservation theorems for limits of structures and global sections of sheaves of structures, *Math. Z.* 166 (1979), 27–54; MR 80d:03029; MR 80d:03029.

[1219] H. Volger, Sheaf representations of algebras and transfer theorems for model-theoretic properties, in *Algebraische Modelle. Kategorien und Gruppoide*, Studien zur Algebra und ihre Anwendungen, Band 7, Akademie–Verlag (1979), 155–163; MR 81m:03050.

[1220] R. F. C. Walters, Sheaves and Cauchy-complete categories, *Cahiers Top. Géom. Diff.* 22 (1981), 283–286; MR 83f:18011.

[1221] R. F. C. Walters, Sheaves on sites as Cauchy-complete categories, *J. Pure Appl. Alg.* 24 (1982), 95–102; MR 83e:18006.

[1222] R. F. C. Walters, *Categories and computer science* (Cambridge Univ. Press, 1993); MR 94i:68001.

[1223] H. Watanabe, A criterion for the existence of subobject classifiers, *Hokkaido Math. J.* 28 (1999), 117–132; MR 99m:18003.

[1224] H. Watanabe, The subobject classifier of the category of functional bisimulations, *Hokkaido Math. J.* 28 (1999), 301–314; MR 00f:18002.

[1225] M. Weinless, The samhita of sets: Maharishi's Vedic science and the foundations of mathematics, *Modern Science and Vedic Science* 1 (1987), 141–204; MR 88i:03016.

[1226] C. Wells, A generalization of the concept of sketch, *Theoret. Comput. Sci.* 70 (1990), 159–178; MR 91d:68082.

[1227] D. Wigner, Two notes on frames, *J. Austral. Math. Soc.* A 28 (1979), 257–268; MR 81d:06016.

[1228] R. Wilkinson, Semigroups in a topos and structure theory of inverse semigroups, *Bull. Austral. Math. Soc.* 26 (1982), 471–472.

[1229] J. T. Wilson, *The assembly tower and some categorical and algebraic aspects of frame theory*, Ph.D. thesis, Carnegie–Mellon Univ. (1994).

[1230] R. J. Wood, $\mathcal{V}$-indexed categories, in *Indexed categories and their applications*, Lecture Notes in Math. vol. 661 (Springer–Verlag, 1978), 126–140; MR 80j:18012.

[1231] R. J. Wood, From tensor products to products, *Commun. Alg.* 8 (1980), 663-669; MR 81i:18008.

[1232] R. J. Wood, Some remarks on total categories, *J. Algebra* 75 (1982), 538–545; MR 83f:18007.

[1233] R. J. Wood, Abstract proarrows I, *Cahiers Top. Géom. Diff.* 23 (1982), 279–290.

[1234] R. J. Wood, Proarrows II, *Cahiers Top. Géom. Diff. Catégoriques* 26 (1985), 135–168; MR 86i:18010.

[1235] G. C. Wraith, Artin glueing, *J. Pure Appl. Alg.* 4 (1974), 345–358; MR 49/9040.

[1236] G. C. Wraith, Lectures on elementary topoi, in *Model theory and topoi*, Lecture Notes in Math. vol. 445 (Springer–Verlag, 1975), 114–206; MR 52/13989.

[1237] G. C. Wraith, Logic from topology: a survey of topoi, *Bull. Inst. Math. Appl.* 12 (1976), 115–119; MR 58/27176.

[1238] G. C. Wraith, Locally equivalent toposes, *Univ. Nat. Zaire Ann. Fac. Sci. Sect. Math.-Phys.* 4 (1978), 155–170; MR 81h:18007.

[1239] G. C. Wraith, Generic Galois theory of local rings, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 739–767; MR 81c:13004.

[1240] G. C. Wraith, Complexification of local fields, in *Topos theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 30 (1979), 211–217; MR 81f:18022.
[1241] G. C. Wraith, Intuitionistic algebra: some recent developments in topos theory, in *Proc. Internat. Cong. Math., Helsinki 1978*, Acad. Sci. Fenn. (1980), vol. I, 331–337; MR 81f:18021.
[1242] G. C. Wraith, Galois theory in a topos, *J. Pure Appl. Alg.* 19 (1980), 401–410; MR 83g:18012.
[1243] G. C. Wraith, Localic groups, *Cahiers Top. Géom. Diff.* 22 (1981), 61–66; MR 82e:18008.
[1244] G. C. Wraith, Toposes and simplicial sets: the cohomological connection, in *Category theoretic methods in geometry*, Aarhus Univ. Var. Publ. Ser. 35 (1983), 281–290; MR 85h:18004.
[1245] G. C. Wraith, Unsurprising results on localic groups, *J. Pure Appl. Alg.* 67 (1990), 95–100; MR 91j:22001.
[1246] G. C. Wraith, Using the generic interval, *Cahiers Top. Géom. Diff. Catégoriques* 34 (1993), 259–266; MR 94j:06018.
[1247] O. Wyler, Are there topoi in topology?, in *Categorical topology*, Lecture Notes in Math. vol. 540 (Springer–Verlag, 1976), 699–719; MR 56/16549.
[1248] O. Wyler, *Lecture notes on topoi and quasitopoi* (World Scientific, 1991); MR 92c:18004.
[1249] O. Wyler, Fuzzy logic and categories of fuzzy sets, in *Non-classical logics and their applications to fuzzy subsets*, Theory Decis. Lib. Ser. B no. 32 (Kluwer, 1995), 235–268; MR 96g:03053.
[1250] A. V. Yakovlev, Galois theory for sheaves of sets (Russian), *Trudy Mat. Inst. Steklov.* 148 (1978), 253–268; MR 81c:12037.
[1251] D. Yetter, Categories of modules in synthetic differential geometry, *Cahiers Top. Géom. Diff. Catégoriques* 28 (1987), 161–170; MR 89d:58007.
[1252] D. Yetter, On right adjoints to exponential functors, *J. Pure Appl. Alg.* 45 (1987), 287–304; MR 88f:18002; corrigenda *ibid.* 58 (1989), 103–105; MR 90c:18001.
[1253] D. Yetter, Models for synthetic supergeometry, *Cahiers Top. Géom. Diff. Catégoriques* 29 (1988), 87–108; MR 89h:58017.
[1254] X.-H. Yuan, Comparisons between category Fuz and category Set, *J. Fuzzy Math.* 5 (1997), 907–914; MR 98h:04013.
[1255] X.-H. Yuan and E. S. Lee, Category Fuz, wtopos, and category FuzFuz, *J. Math. Anal. Appl.* 201 (1996), 489–501; MR 97b:18002.
[1256] D. Zangurashvili, Some categorical-algebraic properties of quasivarieties of algebras in a Grothendieck topos (Georgian), *Soobshch. Akad. Nauk Gruz. SSR* 139 (1990), 25–28; MR 92g:18009.
[1257] A. Zarelua, Sheaf theory and zero-dimensional mappings, in *Applications of sheaves*, Lecture Notes in Math. vol. 753 (Springer–Verlag, 1979), 768–779; MR 81h:55003.
[1258] M. W. Zawadowski, The Skolem–Löwenheim theorem in toposes, *Studia Logica* 42 (1983), 461–475; MR 86b:03085.
[1259] M. W. Zawadowski, The Skolem–Löwenheim theorem in toposes, II, *Studia Logica* 44 (1985), 25–38; MR 87i:03137.
[1260] M. W. Zawadowski, Descent and duality, *Ann. Pure Appl. Logic* 71 (1995), 131–188; MR 96i:03065.
[1261] M. W. Zawadowski, Lax descent theorems for left exact categories, *Dissertationes Math.* (*Rozprawy Mat.*) 346 (1995); MR 96k:03148.
[1262] V. Zöberlein, Doctrines on 2-categories, *Math. Zeitschrift* 148 (1976), 267–279; MR 54/12854.

INDEX OF NOTATION

(*a*) Notation for categories

Notation	Meaning	Introduced
Ab	Category of abelian groups	A1.3.6
$\mathcal{A}_{\mathbb{T}}^{\mathrm{reg}}$	Syntactic allegory of (a regular theory) $\mathbb{T}$	D1.4.15
$\mathbf{Bisep}(\mathcal{C}, J, K)$	Category of biseparated functors on (a bisite) $(\mathcal{C}, J, K)$	C2.2.13
$B_L\mathcal{E}$	Lower bagdomain of $\mathcal{E}$	B4.4.16
Bool	Category of Boolean algebras	D3.4.12
$B_U\mathcal{E}$	Upper bagdomain of $\mathcal{E}$	B4.4.21(*c*)
$[\mathbb{C}]$	Indexed category generated by (an internal category) $\mathbb{C}$	B2.3.3
$\mathbb{C}[A]$	Internal full subcategory of $\mathbb{C}$ generated by A	B2.3.5
$\mathrm{Cart}(\mathbb{C})$	Free cartesian category on $\mathbb{C}$	B4.2.4(*f*)
Cat	Category of small categories and functors	A1.5.1
$\mathcal{C}_B$	Kleisli category of $(-) \times B$	A1.5.4
CBool	Category of complete Boolean algebras	C1.1.21
$\mathcal{C}[\check{\mathcal{E}}]$	Category obtained by splitting idempotents in $\mathcal{E}$	A1.1.8
$\mathcal{C}_{\mathbf{exp}}$	Subcategory of exponentiable objects of $\mathcal{C}$	A1.5.1
Choq	Category of Choquet spaces	A2.6.2(*b*)
CjSLat	Category of complete join-semilattices	C1.1.3
CLat	Category of complete lattices	C1.6.9
$\mathbf{CMon}(\mathcal{C})$	Category of commutative monoids in $\mathcal{C}$	C1.1.8
$\mathrm{CoCart}(\mathbb{C})$	Free cocartesian category on $\mathbb{C}$	B3.2.6
$\mathbf{Cont}(G)$	Category of continuous G-sets	A2.1.6
$\mathbf{Cont}_{\mathcal{S}}(\mathbb{G})$	Category of continuous actions of (a localic groupoid) $\mathbb{G}$	C5.2.11
$\mathbf{Cs}(\mathcal{E})$	Subcategory of coarse objects of (a quasitopos) $\mathcal{E}$	A2.6.11
$\mathcal{C}_{\mathbb{T}}$	Syntactic category of $\mathbb{T}$	D1.4.1
$\mathrm{Dec}_1(\mathbb{C})$	Décalage of $\mathbb{C}$	B2.5.4(*b*)
$\mathbf{\Delta}$	Simplicial category	B2.3.2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

$\mathbb{T}$-**Mod**$(\mathcal{C})$	Category of $\mathbb{T}$-models in $\mathcal{C}$	D1.2.12, D2.1.1(c), D4.3.13
$\mathbb{T}$-**Mod**$(\mathcal{C})_e$	Ditto, with elementary morphisms	D1.2.12
Tors$(\mathbb{C}, \mathcal{E})$	Category of $\mathbb{C}$-torsors in $\mathcal{E}$	B3.2.3
Unif(G)	Category of uniformly continuous G-sets	A2.1.7

(*b*) Notation for 2-categories

Notation	Meaning	Introduced
$\mathfrak{BooLog}$	Boolean toposes and logical functors	D4.3.14(*e*)
$\mathfrak{BooTop}$	Boolean toposes and geometric morphisms	D4.6.12
$\mathfrak{BTop}$	Toposes and bounded geometric morphisms	B3.1.11
$\mathfrak{Cart}$	Small cartesian categories and functors	A4.1.10
$\mathfrak{CART}$	(Possibly) large cartesian categories	B1.2.2(*d*)
$\mathfrak{Cat}$	Small categories and functors	A4.1.4
$\mathfrak{CAT}$	(Possibly) large categories and functors	B1.1.1
$\mathfrak{Cat}_{\mathcal{S}}$	Locally internal categories over $\mathcal{S}$	B2.2.1
$\mathfrak{Cat}(\mathcal{S})$	Internal categories in $\mathcal{S}$	B2.3.1
$\mathfrak{CAT}_{\mathcal{S}}$	$\mathcal{S}$-indexed categories	B1.2.1
$\mathfrak{Cau}_f$	Finite Cauchy-complete categories	C2.2.23
$\mathfrak{cFib}_{\mathcal{S}}$	Cloven fibrations over $\mathcal{S}$	B1.3.6
$\mathfrak{Cocomp}_{\mathcal{S}}$	Cocomplete locally small $\mathcal{S}$-indexed categories	B4.5.1
$\mathfrak{Coh}$	(Small) coherent categories and coherent functors	D3.5.5
$\mathfrak{Comp}_{\mathcal{S}}$	Complete locally small $\mathcal{S}$-indexed categories	B4.5.3
$\mathfrak{DMTop}_{sk}$	De Morgan toposes and skeletal maps	D4.6.12
$\mathfrak{DTop}/\mathcal{S}$	Diagram toposes over $\mathcal{S}$	B3.2.8(*c*)
$\mathfrak{Fib}_{\mathcal{S}}$	Fibrations over $\mathcal{S}$	B1.3.6
$\mathfrak{FILT}_{\mathcal{S}}$	$\mathcal{S}$-indexed categories with filtered colimits	C4.3.3
$\mathfrak{GFrm}_{\mathcal{S}}$	Giraud frames over $\mathcal{S}$	B4.5.1
$\mathfrak{LocPres}_{\mathcal{S}}$	Locally presentable $\mathcal{S}$-indexed categories	B4.5.2
$\mathfrak{LocTop}/\mathcal{S}$	Local toposes over $\mathcal{S}$	C3.6.14
$\mathfrak{Log}$	Toposes and logical functors	D4.3.13
$\mathfrak{Log}_N$	Ditto, with natural number objects	D4.3.19(*d*)
$\mathfrak{LTop}$	Toposes and localic morphisms	A4.6.12
$\mathfrak{PreTop}$	(Small) pretoposes and coherent functors	D2.2.5
$\mathfrak{Prof}_{\mathcal{S}}$	Internal profunctors in $\mathcal{S}$	B2.7.1
$\mathfrak{pTop}/\mathcal{S}$	Pointed bounded $\mathcal{S}$-toposes	C3.6.13
$\mathfrak{PTop}/\mathcal{S}$	Ditto (lax version)	C3.6.14
$\mathfrak{Reg}$	(Small) regular categories and regular functors	D2.2.3
$\mathfrak{rTors}_{\mathcal{S}}$	Profunctors which are right torsors	B3.2.8(*c*)
$\mathfrak{sFib}_{\mathcal{S}}$	Split fibrations over $\mathcal{S}$	B1.3.6
$\mathfrak{Span}(\mathcal{C})$	Bicategory of spans in $\mathcal{C}$	A3.3.8

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

(c) Miscellaneous notation

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

GENERAL INDEX

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2

Sections A and B are in Vol. 1; Sections C and D are in Vol. 2